DOCUMENT RESUME

ED 233 076 TM 830 533

AUTHOR Wilson, Kenneth
TITLE A Review of Research on the Prediction of Academic Performance after the Freshman Year.
INSTITUTION College Entrance Examination Board, New York, N.Y.
REPORT NO CB-R-83-2; ETS-RR-83-11
PUB DATE 83
NOTE 60p.; Some tables and figures may be marginally legible because of small print.
AVAILABLE FROM College Board Publications, Box 886, New York, NY 10101 ($5.00).
PUB TYPE Information Analyses (070)

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS *Academic Achievement; *College Entrance Examinations; College Freshmen; Comparative Analysis; *Grade Point Average; Higher Education; Performance Factors; *Predictive Measurement; Predictive Validity; *Undergraduate Students

ABSTRACT

The criterion most frequently used in studies designed to assess the predictive validity of measures used in college admission has been the freshman-year grade point average (GPA). It is not self-evident that the first-year GPA provides either a sufficient or a representative sample of a student's academic performance. Questions have been raised regarding the validity of admissions measures for predicting longer-term performance in college. This is the report of a systematic review of research bearing on (a) the validity of admissions measures for predicting GPA that reflects performance beyond the freshman year--i.e., longer-term cumulative or independently computed post-freshman-year GPA, such as senior-year GPA; (b) the comparative relevance and utility of freshman-year, cumulative, and independently computed post-freshman-year GPA as criteria for the validation of admissions measures. Among other things, the research reviewed lends support to the traditional practice of employing the freshman-year GPA in admissions-related predictive validity studies. (Author)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED233076

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

College Board Report

No. 83-2

A Review of Research on the Prediction of Academic Performance After the Freshman Year

Kenneth M. Wilson

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

S. W. Gardner

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

A Review of Research on the Prediction of Academic Performance After the Freshman Year

Kenneth M. Wilson

Educational Testing Service

College Board Report No. 83-2

ETS RR No. 83-11

College Entrance Examination Board, New York, 1983

To the College Board whose support made this review possible; to Frances Livingston, Alice Norby, and Christine Sansone and her associates for assistance in preparing the manuscript; to Thomas Donlon and Donald Powers for critical reviews of earlier versions of the manuscript; these contributions are acknowledged with appreciation.

Researchers are encouraged to express freely their professional judgment. Therefore, points of view or opinions stated in College Board Reports do not necessarily represent official College Board position or policy.

The College Board is a nonprofit membership organization that provides tests and other educational services for students, schools, and colleges. The membership is composed of more than 2,500 colleges, schools, school systems, and education associations. Representatives of the members serve on the Board of Trustees and advisory councils and committees that consider the programs of the College Board and participate in the determination of its policies and activities.

Additional copies of this report may be obtained from College Board Publications, Box 886, New York, New York 10101. The price is $5.

Copyright © 1983 by College Entrance Examination Board.
All rights reserved.
Printed in the United States of America.

CONTENTS

List of Tables iii

List of Figures iv

List of Exhibits iv

Abstract 1

Introduction 1

Predicting Post-freshman-year GPA: Some Generalizations Based on a Review of Research Findings . . . 2

Some Working Hypotheses Regarding Post-freshman-year Prediction 2

Predicting Post-freshman-year GPA: A Look at the Research Record 3

Predicting the General Cumulative GPA 3
Predicting Independent Post-freshman-year GPA 13
Predicting Later Grades from Earlier Grades 26
Predicting Later GPA from Earlier GPA and Admissions Measures 32

Recapitulation 36

Implications for Validation Research 37

Implications for the Use of the Freshman GPA in Validity Studies 38
General Implications for Validation Research 38

References 40

Appendix 43

Tables

1. Studies Involving the Prediction of Cumulative GPA for Periods Exceeding the Freshman Year, with or without Comparative Coefficients for Shorter-term GPA and/or Other Criteria 7

2. Distributions of Shrunken Multiple Correlation Coefficients for Selected Elements of the Washington Pre-College Test Battery vs. Year 1 and Four-year Cumulative GPA, by Subject 12

3. Correlation between Repeated Measures in a Series: Initial vs. Final Measure, First vs. Second Measure, and Last Two Measures in the Series 17

4. Correlational Data on the Stability of Standing on Admissions Tests Over Various Time Intervals 22

5. Data Indicating a Declining Gradient in the Correlation of Various Admissions Measures with Independently Computed GPA in Successive Undergraduate Semesters in Restricted Samples Including Only Seniors or Graduates 25

6. Correlations between GPA in Adjacent Undergraduate Semesters in Restricted Samples Including Only College Seniors or Graduates 25

7. Declining Gradients in Validity of Year 1 and Semester 1 GPA as Predictors of GPA in Successive Subsequent Undergraduate Semesters in Samples Including Seniors or Graduates Only 27

8. Studies Involving the Prediction of Independently Computed Post-freshman-year GPA Using both Earlier College GPA and Admissions Measures as Predictors 30

9. Results of Special Reanalyses Indicating the Declining Incremental Contribution of Admissions Measures to the Prediction of a Remote Semester GPA Criterion as GPA in Successive Earlier Semesters Are Introduced as Predictors 34

10. Results of Special Reanalyses Indicating the Slight Contribution of Admissions Measures to Prediction of GPA in Successive Semesters after Taking into Account the Previous Semester GPA Record 34

Figures

1. Trends in mean semester GPA for students classified by number of consecutive semesters of attendance: Fall 1958 entering freshmen, University of Wisconsin 5

2. Trends in mean semester GPA for students classified by number of consecutive semesters of attendance: Fall 1971 entering freshmen, anonymous university 6

3. Trends in the relationship of entry-level test scores to independently computed GPA in successive years, semesters, or quarters, and of earlier to later undergraduate GPA 15

4. Gradients in correlation with GPA by semester of tests administered near the beginning and near the end of college 19

5. Predictive and postdictive correlations with GPA in successive quarters 20

6. Trends in the correlation of GPA in successive undergraduate semesters with (a) all previous semester GPA, (b) GPA in the immediately preceding semester only, (c) Year 1 GPA, (d) Semester 1 GPA, and (e) a regression-weighted composite of available admissions measures, based on ad hoc supplementary analyses of correlation matrices, shown in the Appendix, for samples of seniors or graduates from the University of Illinois, anonymous university, and selective college 29

Exhibits

A. Intercorrelations, means, and standard deviations of GPAs in 12 curricular areas, high school GPA, and SATs, and distributions of correlations between the admissions variables and the curricular GPA 14

B. Correlational data exhibiting a declining gradient in relationships between repeated measures in a series as the time interval between measures increases 16

C. Studies involving the prediction of college GPA from GPA in elementary grades and from very early test performance 23

D. The relationship of cumulative GPA at the end of successive quarters to later cumulative and senior-year GPA, and of first-quarter GPA to GPA in successive quarters, with the corresponding intercorrelation matrix for GPA in each of 12 successive quarters 28

E. The relationship between SAT scores, high school average (HSA), Predicted Grade (PG), and freshman average, and various measures of performance 33

ABSTRACT

The criterion most frequently used in studies designed to assess the predictive validity of measures used in college admission has been the freshman-year GPA. It is not self-evident that the first-year GPA provides either a sufficient or a representative sample of a student's academic performance. Questions have been raised regarding the validity of admissions measures for predicting longer-term performance in college. This is the report of a systematic review of research bearing on (a) the validity of admissions measures for predicting GPA that reflects performance beyond the freshman year—i.e., longer-term cumulative or independently computed post-freshman-year GPA, such as senior-year GPA; (b) the comparative relevance and utility of freshman-year, cumulative, and independently computed post-freshman-year GPA as criteria for the validation of admissions measures. Among other things, the research reviewed lends support to the traditional practice of employing the freshman-year GPA in admissions-related predictive validity studies.

INTRODUCTION

It is axiomatic that measures used to assess the academic qualifications of applicants for admission to educational programs should have demonstrated validity for predicting relevant measures (criteria) of academic performance in those programs. To assess the predictive, or criterion-related, validity of preadmissions measures, literally thousands of college prediction studies have been conducted. More or less standard in design and methodology, these widely replicated studies typically have called for analysis of the relationship of selected criterion measures (almost always the freshman grade point average, or GPA) to available admissions measures (typically, scores on standard admissions or other tests of developed cognitive abilities or achievement, and the high school record) to answer questions such as the following:

- How valid are the admissions measures, individually and in best-weighted composites, for predicting freshman GPA, and what is their relative contribution to the best-weighted composite predictor?

The practical contribution of these basic validity studies has been (a) to facilitate admissions by providing, through multiple regression equations, a systematic basis for estimating the probable academic performance of prospective students, in the form of a predicted freshman GPA, and (b) to demonstrate test validity in an "accountability" context.

Results of numerous first-year validity studies have been summarized periodically over the past half century (e.g., Segel, 1934; Fishman and Pasanella, 1960; Fishman, 1963; Schrader, 1971; ACTP, 1973; Fincher, 1974; Ramist, draft). Thus, a substantial body of empirical evidence is available on an important aspect of the validity of the most commonly used admissions measures—the high school record and scores on tests such as the College Board Scholastic Aptitude Test (SAT) and the American College Testing Program battery (ACT)—namely, their validity for predicting freshman-year performance in college.

Over the same time period, however, among those concerned with selective admission remarkably little interest has been shown in corresponding questions regarding another aspect of the validity of admissions measures, namely, their validity for predicting GPA criteria reflecting performance *after* the freshman year. For example, few investigators appear to have been interested in using either the overall cumulative GPA, which partially reflects post-freshman-year performance, or a GPA based exclusively on work completed after the freshman year (e.g., a general post-freshman-year GPA, upper-division GPA, senior-year GPA) as criteria in predictive validity studies.

The relevance of a student's GPA, per se, as an academic performance criterion in admissions-related predictive validity studies is essentially self-evident, given (a) the centrality of academic course work in colleges, (b) the traditional acceptance of grades as the official measure of a student's performance level in course work, and (c) the fact that students must maintain a specified minimum GPA during each marking period and cumulatively to avoid academic probation or dismissal and to continue toward program completion and graduation. It is not self-evident, however, that the first-year performance record alone represents either a sufficient or a representative sample of a student's academic performance for the purpose of assessing the predictive validity of admissions measures.

Are admissions measures as valid for predicting the overall cumulative record or, say, the senior-year GPA, as they are for predicting the freshman-year GPA? Should they be expected to be? How may differences in validity, if any, be accounted for? Can the traditional practice of using the freshman-year GPA in admissions-related predictive validity studies be justified? These are reasonable questions that the results of numerous replications of first-year studies do not answer. Generalizations about the validity of admissions measures for predicting GPA criteria that partially or exclusively reflect performance beyond the freshman year require evidence based on studies using those post-freshman-year criteria.

This report presents the results of a systematic effort to locate, summarize, and evaluate available research evidence bearing on (a) the validity of precollege admissions measures for predicting cumulative and independently computed post-freshman-year GPA criteria, and (b) the comparative relevance and utility of freshman-year, cumulative, and post-freshman-year GPA as criteria for assessing the predictive validity of admissions measures.

In the search for relevant studies, several procedures were followed:

- Reviews of the literature on prediction, such as those cited earlier, were checked.
- The tables of contents of several relevant journals (generally for the period 1965–1981) were examined (e.g., *Educational and Psychological Measurement*, *Journal of Educational Psychology*, *American Educational Research Journal*, etc.).
- Guides to periodical literature were consulted.
- A computerized search of ERIC data bases was conducted.
- Leads suggested by references in published and unpublished reports were followed to the maximum extent possible.

Based on these procedures, the research reviewed here is believed to be representative of the available body of evidence bearing on the validity of admissions measures for predicting post-freshman-year GPA, generally, as well as on the comparative predictability, and relevance for admissions-related validation research, of freshman-year, cumulative, and post-freshman-year GPA criteria.

PREDICTING POST-FRESHMAN-YEAR GPA: SOME GENERALIZATIONS BASED ON A REVIEW OF RESEARCH FINDINGS

The studies located differed widely in objectives, design, scope, and combinations of predictor-criterion variables with little or no replication. They covered a span of some 60 years. None of them was concerned with, or provided evidence regarding, all of the principal themes or issues under consideration. Longitudinal studies providing evidence regarding the relationship of preadmissions predictors to independently computed GPAs in successive marking periods (quarters, semesters, years) and cumulatively from the freshman through the senior year, as well the interrelationships among the GPA variables themselves, have rarely been conducted (or reported)—only a handful were located in this search.

Based on the literature search, questions regarding the prediction of academic performance beyond the freshman year appear to have been of little interest to those concerned with selective admission. Comparatively few studies were located that were concerned either directly or incidentally with one or more of the following interrelated issues or concerns, none of which appears to have been the focus of sustained inquiry.

1. The validity of standard admissions or other tests of developed cognitive abilities or achievement, and/or the high school record for predicting:
   a. the cumulative GPA as opposed to the freshman-year GPA;
   b. one or more independently computed post-freshman-year GPAs.
2. The validity of the freshman-year GPA for predicting post-freshman-year GPA, or, more generally, the stability of the relative standing of students in terms of general GPA across successive marking periods from the freshman year through the senior year.
3. The incremental contribution of standard admissions or other tests of cognitive abilities and/or the high school record to the prediction of post-freshman-year GPA after taking into account the freshman-year GPA.

Some Working Hypotheses Regarding Post-freshman-year Prediction

Despite these limitations, certain trends in the findings are sufficiently clear to warrant formulating some generalizations or working hypotheses. These generalizations are presented below, without specific documentation, as the framework for a more detailed consideration of the related evidence, its limitations, and its implications (a) for the practice of using freshman-year GPA in validity studies, and (b) for validation research generally.

Validity of Preadmissions Measures for Predicting the Cumulative GPA

Preadmissions measures (such as commonly used standard admissions tests and the high school record) are essentially as valid for predicting the overall (or long-term) cumulative record as for predicting the short-term, freshman GPA criterion traditionally used in validity studies. The correlation between admissions measures and GPA tends to remain relatively stable as the grade record accumulates from the freshman through the senior year. Composite-predictors that are best-weighted for predicting freshman GPA are likely to correlate about as well with longer-term cumulative GPA, and vice versa.

Validity of Preadmissions Measures for Predicting Post-freshman-year GPA

Preadmissions measures tend to be less valid for predicting post-freshman-year GPA than for predicting either the cumulative or the freshman-year GPA. More generally, the correlation between admissions measures and independently computed GPA in successive marking periods (quarters, semesters, years) tends to exhibit a gradually declining gradient from the freshman through the senior year. This declining validity gradient appears to hold for the high school record and for standard admissions tests individually, as well as for best-weighted composite predictors.

- The pattern of declining validity does not appear to be due primarily either to statistical artifacts, such as restriction of range on the admissions measures, decrease in GPA standard deviation, or reliability, or to major changes over time in the rank ordering of students in the abilities measured by standard admissions tests—e.g., admissions tests administered toward the end of the college senior year also tend

to be related more closely to freshman-year GPA than to the college senior-year GPA.

Validity of Freshman-year GPA for Predicting Later GPA

Although the freshman-year GPA tends to be more valid than preadmissions measures for predicting post-freshman-year GPA, the correlation of freshman-year GPA with GPA in successive marking periods shows the same generally declining gradient from the freshman through the senior year as that for preadmissions measures, but at a generally higher level. Considerable change in rank ordering with respect to GPA takes place between the freshman and the senior year in samples of "successful" college students.

- Although the level of correlation between freshman-year GPA and GPA in later terms declines as the time interval between terms increases, the level of correlation between GPA in successive adjacent terms appears to remain relatively stable, as does the level of correlation between the general cumulative GPA at the end of any given term and (a) GPA in the immediately following term and, (b) the cumulative GPA for all subsequent terms.

Contribution of Preadmissions Measures to Predicting Post-freshman-year GPA after Taking into Account the Freshman-year GPA

Commonly used preadmissions measures have little or no incremental validity when combined with freshman-year GPA to predict later GPA within a given setting. The best prediction of GPA in any given marking period is provided by the cumulative GPA prior to that marking period.

- Although preadmissions measures tend to be correlated modestly with all post-freshman-year GPA criteria, their contribution to the prediction of general post-freshman-year GPA is indirect, through the post-freshman-year GPA-related variance that they have in common with the freshman-year GPA.

PREDICTING POST-FRESHMAN-YEAR GPA: A LOOK AT THE RESEARCH RECORD

Research evidence for the several tentative generalizations and working hypotheses about post-freshman-year prediction outlined above is reviewed in the following sections.

Predicting the General Cumulative GPA[1]

A limited number of studies have provided empirical evidence bearing on the proposition that standard admission tests and other tests of developed cognitive abilities or achievement, and the high school record, are essentially as valid for predicting the longer-term cumulative GPA as for predicting the freshman-year GPA. Before looking at the empirical evidence for this proposition, it is useful to examine some general considerations that should be kept in mind in evaluating and interpreting the findings of studies involving prediction of the general cumulative GPA alone or the comparative prediction of freshman-year and cumulative GPA criteria.

Some General Interpretive Considerations

It is useful to recognize, as noted elsewhere (Wilson, 1981), that admissions measures should be expected to correlate about as well with longer-term cumulative GPA as with the freshman-year GPA on purely statistical grounds, *because the freshman-year GPA is a major continuing source of variance in any longer-term cumulative GPA*. Snedecor (1956, p. 189), for example, has shown that the correlation of a part (A), with a whole (A + B) will be high (r = .707), even if A and B are completely uncorrelated (r = .00), in the special case when A and B have equal variances. Actual freshman-year/cumulative GPA correlations in the .8 range appear to be typical in undergraduate samples (see for example, Jones and Case, 1955; Willingham, 1962; Wilson, 1967, 1978, 1980).[2] Given such built-in criterion stability as the GPA accumulates, the freshman-year GPA and the cumulative GPA would appear to be statistically interchangeable for the purpose of assessing the predictive validity of admissions measures.

A second consideration to be kept in mind in evaluating the empirical findings is that two types of cumulative GPA have been employed in validity studies. Some investigators have used a *fixed-base cumulative GPA* that can be computed only for students who complete a specified number of years of study (e.g., two, three, or four years) while others have used a *variable-base cumulative GPA* that can be computed for essentially all members of an entering cohort because it is based on all course work completed *regardless of amount* during a given study period—for example, in a four-year study, the variable-base cumulative could be computed for all students who completed at least one term (quarter or semester).

Any study using a fixed-based cumulative GPA would necessarily have to be based on a selected group of persisting students who are likely to be a biased sample of the entering cohort, having been explicitly selected in terms of previous grades and secondarily in terms of correlates of the earlier grades (e.g., the admissions measures). On the other hand, studies employing a variable-base cumulative GPA can include essentially all members of an entering freshman

[1] The term "cumulative GPA" ordinarily refers to any GPA based on grades accumulated over two or more official marking periods, usually quarters or semesters. Thus, the freshman-year GPA is a "short-term" cumulative GPA. The intended reference in this review is to longer-term cumulative GPA, including grades for a period of study longer than the freshman year.

[2] This level of correlation between Year 1 and overall cumulative GPA has also been reported for professional school samples—for example, law school, .84 (Carlson and Werts, 1976), medical school, .82, median for ten successive classes (Gough, Hall, and Harris, 1964).

cohort, regardless of the period of time covered by the study. Thus, in a four-year study employing a variable-base cumulative GPA, the GPAs involved could be based on, say, from one semester to four years of course work, but the entire range of performance on the criterion and predictor variables would be represented in the study sample. However, if the four-year cumulative GPA were to be used, a restricted range of performance on both the GPA criterion and the predictor variables would be represented in the sample. Theoretically, such differences have implications for the interpretation of observed validity coefficients.

Some indication of the potential differences in the range of performance that might be represented in a four-year validity study as a function of the type of cumulative GPA employed may be seen in Figures 1 and 2. These figures show trends in mean semester GPA for students, in two state universities, classified by number of consecutive semesters of attendance.

Trends are generally similar in both settings: persisting members of the entering cohort had substantially better initial mean GPA than their nonpersisting counterparts, and the GPA gap between persisters and nonpersisters tended to increase semester by semester. It is evident that in a four-year study, use of the fixed-base, four-year cumulative GPA would result in the elimination of " . . . the range of test scores and grades that is of most interest to admissions officers . . . " (French, 1958). One result, theoretically, would be attenuation of the relationship between the admissions variables and the fixed-base cumulative GPA due to restriction on the GPA criterion, a factor which would not be involved in a four-year study employing a variable-base cumulative GPA.

In a closely related vein, a third consideration is that in some studies of the comparative validity of admissions variables for predicting first-year and fixed-base cumulative GPA, validity coefficients involving the freshman-year GPA may be based on all entering freshman, whereas those involving the cumulative GPA may be based only on individuals completing two, three, or four years of study—that is, some comparisons are based on samples with a constant N (e.g., only graduating students are used for computing both first-year and cumulative coefficients) whereas others are based on a changing N (e.g., all freshman vs. all graduating seniors). Under changing N conditions, range restriction might be expected to result in artifactually lower validities when the long-term cumulative GPA is used than when the first-year criterion is used.

The Empirical Record

Very few studies involving prediction of cumulative GPA criteria were located in the review. Table 1 includes for each of 21 studies a brief description of the study context and sample, the predictor and criterion variables employed, and salient findings. The studies differed widely in design, combinations of predictor and criterion variables, and scope, and they were characterized by essential lack of replication. Six provided validity data for a cumulative criterion only. Accordingly, detailed results are not amenable to succinct summarization, and the studies do not lend themselves to classification in terms of methodology, types of criteria, duration of study period, or other theoretically relevant variables.

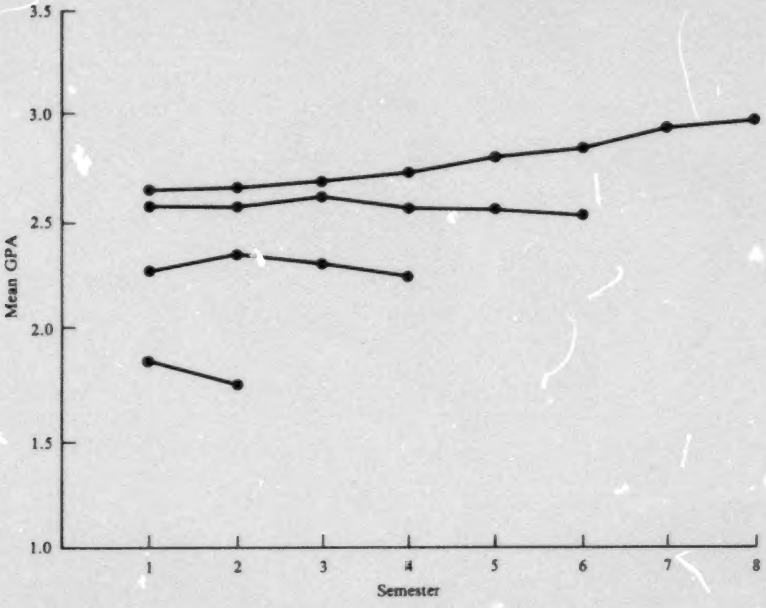
Study descriptions are ordered chronologically from earlier to more recent dates. As may be determined by inspection, the studies are rather evenly distributed chronologically—seven were conducted prior to 1960, six during the 1960s, and eight after 1970. Since its inception in 1964, the College Board Validity Study Service at ETS, has alone conducted over 3,000 validity studies for more than 1,000 undergraduate institutions, essentially all of which involved first-year GPA criteria. Given such volume for first-year studies, the paucity of studies involving longer-term GPA criteria attests to a general lack of interest in post-freshman-year prediction. Several of the longer-term post-1970 studies were prompted by an interest in questions regarding the use of standard admissions tests with minority students (e.g., Farver, Sedlacek, and Brooks, 1975; Wilson, 1978, 1980, 1981).[3]

It is noteworthy that empirical support for the proposition that admissions variables are essentially as valid for predicting longer-term cumulative GPA as for predicting shorter-term GPA is provided by studies differing with respect to the sampling considerations outlined above that theoretically might be expected to have some effect on study outcomes.

- For example, in a major multiinstitutional study involving eight private and two public east coast colleges (French 1957, 1958), basic admissions variables (SAT-V, SAT-M, and HSR) were found to be essentially as valid for predicting a variable-base, final general cumulative GPA as for predicting the freshman-year GPA (using in both instances the entire range of predictor/criterion variance in the entering cohort in a constant N sample). In another multiinstitutional study involving 31 sets of data for 16 colleges and over 3,000 students in University System of Georgia institutions, Hills, Bush, and Klock (1964), using changing N samples and fixed-base cumulative GPA criteria, found that the same set of admissions variables could be used to predict the four-year average at graduation and the cumulative sophomore average (with restricted ranges of predictor/criterion variance) as well as the freshman-year GPA (representing the entire range of performance), with correlations averaging in the .60+ range. Olsen and Schrader (1959), in a third major multiinstitutional study, used two-year and four-year cumulative GPA (fixed-base) criteria, also for changing N samples, in nine colleges and concluded that for SAT candidates tested twice (about a year apart), preliminary and final SAT scores had virtu-

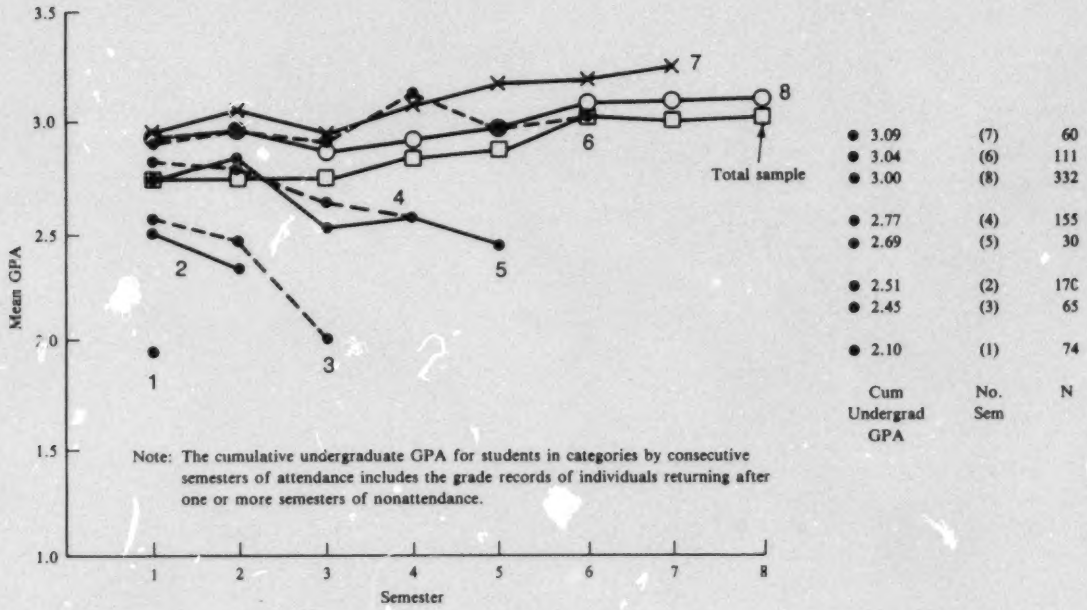
[3]See Breland (1978) for some additional studies involving minority students and longer-term GPA criteria.

Figure 1. Trends in mean semester GPA for students classified by number of consecutive semesters of attendance: Fall 1958 entering freshmen, University of Wisconsin



Source. Lins and Abell (1965)

Figure 2. Trends in mean semester GPA for students classified by number of consecutive semesters of attendance: Fall 1971 entering freshmen, anonymous university



Source. Wilson (1978, 1980) Unpublished findings

Table 1. Studies Involving the Prediction of Cumulative GPA for Periods Exceeding the Freshman Year, with or without Comparative Coefficients for Shorter-term GPA and/or Other Criteria (e.g., major-field GPA, graduation/nongraduation)

Source	*Variables*[a]	*Salient Findings*[b]

Dwyer (1940), cited in French (1957)

Cohorts entering Univ of Michigan, 1927–1936; liberal arts men and women, and engineers (internal study)

SAT*, HSR
Iowa English
Iowa Math
Yr 1 GPA, Cum GPA
*(SAT not further described; may be sum or weighted composite of V and M)

	Men (Liberal Arts)				Women (Liberal Arts)				Engineers			
Predictor	Yr 1		Cum		Yr 1		Cum		Yr 1		N	Cum
SAT	(536)*	.42	(207)*	.37	(285)*	.49	(150)*	.52	(187)*	.35	(72)	.22
HSR	(595)	.43	(225)	.51	(298)	.39	(153)	.39	(203)	.38	(73)	.49
Math	(541)	.31	(213)	.28	(280)	.30	(50)	.41	(190)	.38	(72)	.37

*(Ns–note reduction in sample size for cumulative GPA.)

Brush (1941), cited in French (1957)

Engineers, setting not identified

Thorndike Intelligence Test
Columbia Physics
Columbia Geometry
Other tests
First-semester GPA
Cumulative GPA

Predictor	First-term GPA	Cumulative GPA	
Thorndike	.47	.43	(N = 104, constant)
Columbia Physics	.59	.50	
Columbia Geometry	.32	.38	

Jones and Case (1955)

Junior-level engineering students at UCLA (N = 199)

HSR
Technical Vocabulary
Quantitative Inference
Spatial Visualization
Yr 1 GPA, Yr 2 GPA

Predictor	Yr 1 GPA	Yr 2 GPA	2-yr Cum GPA
HSR	.36	.33	.38
Technical Vocab	.35	.31	.36
Quant Inf	.20	.21	.23
Spatial Visual	.30	.20	.26
Yr 1 GPA	–	.66	.86
Yr 2 GPA		–	.92

University (1956) Unpublished report, cited in French (1957)

Large private coed university

SAT-V
SAT-M
HSR
Cumulative GPA
Major-field GPA

	GPA in major field							
Predictor	Gen Cum	Soc Sci	Educ	Humanities	Bio Sci	Phys Sci	Engineering	(Mean)*
HSR	.54	.53	.50	.53	.46	.71	.63	.548
SAT-V	.46	.48	.38	.41	.49	.50	.40	.457
SAT-M	.22	.16	.24	.33	.46	.56	.34	.282
N	(957)	(448)	(59)	(143)	(111)	(86)	(110)	(957)

*Note that these weighted averages of major-field coefficients (pooled within-major correlations) and coefficients for the general Cum GPA without regard to major are very similar.

Frederiksen (undated), cited in French (1957)

Engineers, Princeton University (internal study)

SAT-Verbal
SAT-Mathematical
HSR
CB Chem Achievement
Other achievement measures
First-semester GPA
Cumulative GPA

Predictor	First-term GPA		Cumulative GPA		
SAT-V	(516)*	.42	(131)*	.49	(*Ns–note reduction in sample size for cumulative GPA)
SAT-M	(236)	.33	(58)	.17	
HSR	(499)	.54	(124)	.53	
CB Chem	(204)	.54	(52)	.59	

Table 1 (cont.). Studies Involving the Prediction of Cumulative GPA for Periods Exceeding the Freshman Year, with or without Comparative Coefficients for Shorter-term GPA and/or Other Criteria (e.g., major-field GPA, graduation/nongraduation)

Source	*Variables*[a]	*Salient Findings*[b]

French (1957, 1958)

Eight private, two public east coast colleges; nine liberal arts, one technical; 1951 entering cohorts; range in selectivity from mean SAT-V = 637, SD = 64 to SAT-V = 452, SD = 100.

SAT-Verbal
SAT-Mathematical
HSR
Experimental College Board battery
Yr 1 GPA
Final Cum GPA based on all courses completed
GPA by course areas
Yr 1 GPA
Yr 3–4 GPA (Major)
Graduation/nongraduation

				GPA by course areas					
	General GPA			Science		Soc Sci		Humanities	
Predictor	Yr 1	Cum	Gr/NonGr*	Yr 1	Major	Yr 1	Major	Yr 1	Major
SAT-V	.44	.43	.09	.36	.35	.43	.43	.39	.34
SAT-M	.29	.27	.06	.37	.34	.20	.26	.18	.23
HSR	.46	.46	.18	.41	.27	.34	.29	.37	.29
Cum GPA	n.a.	–	.44						

Note: Coefficients are weighted averages; for general GPA Yr 1 vs. Cum GPA analyses Ns were approximately 3,500; for analyses of GPA by course areas, Yr 1 Ns are larger than Ns for major GPA analyses which are based on GPA in field during junior and senior years only.

*Percent graduating ranged from 82 through 47; for percent graduating vs. SAT-V mean over 10 colleges, rho = .53.

Olsen and Schrader (1959)

Three men's colleges, three women's colleges, and three engineering schools; students taking SAT twice (March or May 1953 and January or March 1954); selectivity ranged from mean SAT-V (final) = 483, SD = 81 through mean = 608, SD = 61

Preliminary SAT-V
Preliminary SAT-M
Final SAT-V
Final SAT-M
2-yr Cum GPA
4-yr Cum GPA

College		A	B	C	D	E	F	G	H	I
		Men			Women			Engineering		
Predictor		Multiple correlation for SAT-V plus SAT-M								
Prelim	2-yr cum	.46	.46	.51	.46	.42	.36	.36	.26	.32
SAT	4-yr cum	.51	.33	.47	.49	.40	.36	.47	.36	.22
Final	2-yr cum	.47	.43	.53	.48	.40	.42	.35	.32	.27
SAT	4-yr cum	.50	.30	.50	.51	.37	.36	.35	.33	.29
N	2-yr cum	239	253	134	369	262	139	116	85	54
	4-yr cum	205	211	99	291	216	111	89	48	45

In Prelim analyses, validities higher for 4-yr than 2-yr cum in four samples, lower in four and equal in one.

In Final analyses, validities higher for 4-yr than for 2-yr cum in four samples, lower in four and equal in one.

Chase (1960)

Former students at Hunter College, 1936–1956 who earned doctorates by Fall, 1957 (N = 307)

HSR
Term 1 GPA
Cumulative GPA (4-yr)

Predictor	Cum GPA
HSR	.66
Term 1 GPA	.86*

N = 307

*(In a random sample of students, r = .71)

Willingham (1962)

Students who completed more than three quarters at Georgia Institute of Technology (N = 799)

HSR
SAT-V
SAT-M
Predicted Yr 1 GPA
Yr 1 GPA
Yr 2 GPA
Final cumulative GPA

Predictor	Yr 1 GPA	Yr 2 GPA	Final Cum GPA
HSR	.48	.36	.45
SAT-V	.45	.27	.36
SAT-M	.45	.29	.37
Pred Yr 1 GPA	.63	.43	.55
Year 1 GPA	–	.63	.83
Year 2 GPA		–	.84

Table 1 (cont.). Studies Involving the Prediction of Cumulative GPA for Periods Exceeding the Freshman Year, with or without Comparative Coefficients for Shorter-term GPA and/or Other Criteria (e.g., major-field GPA, graduation/nongraduation)

Source	*Variables*[a]	*Salient Findings*[b]
Hills, Bush, and Klock (1964) Units of the University System of Georgia, analysis of 31 sets of data for 16 colleges and 3,303 individual students. Samples less highly selected than, for example, those studied by French (1958) or Olsen and Schrader (1959)	SAT-V SAT-M HSR Predicted Yr 1 GPA Yr 2 GPA Cum 2-yr GPA Cum 4-yr GPA	Males: Yr 2 GPA; 2-yr Cum; 4-yr Cum — Females: Yr 2 GPA; 2-yr Cum; 4-yr Cum V, M, HSR (R)* — .58 .67 — .68 .65 Predicted Yr 1(r)** .52 .60 .63 .67 .66 .61 *No. studies (R) — 6 2 — 6 3 Total N (R) — 666 267 — 289 271 **No. studies (r) 9 7 2 9 6 2 Total N (r) 650 432 239 685 257 204 Range of correlations = .52–.68, median = .64; average V, M, HSR multiple vs. Year 1 GPA in the System reported to be R = .66. *Note:* Coefficients are average values across studies.
Sgan (1964) Two graduating classes, Brandeis University, 1962 and 1963; students with Latin Honors (highest GPA at graduation and an equal number of students with lowest GPA in class)	SAT-V SAT-M Cumulative GPA (4-yr)	Graduates with honors (highest cum GPA): 1962, 1963 — Equal number with lowest cum GPAs: 1962, 1963 SAT-V vs Cum GPA (r) .40 .27 .40 .06 SAT-M vs Cum GPA (r) .38 .18 .37 .15 SAT-V (median) 674 669 567 571 No. in subgroup (60) (54) (60) (54)
Prien and Botwin (1966) Small private women's college	SAT-V, SAT-M Yr 1 GPA Yr 2 GPA Sem 1 of Yr 3 GPA 2-yr Cum GPA	Yr 1 — Yr 2 — Sem 1, Yr 3 — 2-yr Cum SAT-V .48 .41 .35 .46 SAT-M .41 .44 .32 .45 (N = 95, constant)
Wilson (1967) Graduating seniors, selective women's college	HSR SAT-V SAT-M CB Achievement Average Sem 1 GPA, Yr 1 GPA, Yr 3 GPA, Yr 4 GPA, and 4-yr Cum GPA	Predictor — Sem 1 — Yr 1 — Yr 3 — Yr 4 — 4-yr Cum HSR .36 .38 .27 .23 .35 SAT-V .08 .12 .13 .09 .18 SAT-M .23 .23 .11 .08 .16 N = 259, constant CB Ach .31 .34 .25 .16 .34 SAT-V Mean = 575 Sem 1 GPA .92 .50 .45 .74 SD = 65 Yr 1 GPA .55 .49 .81 Yr 3 GPA .70 .85 Yr 4 GPA .81 SD .51 .47 .42 .43 .37
Mauger and Kolmodin (1975) Small denominational liberal arts college; selectivity indicated by mean SAT-V = 494, SD = 85; two samples—one entering cohort studied longitudinally, the other comprised of graduates only	SAT-V SAT-M Yr 1 GPA Final cumulative GPA Other variables	Predictor — Yr 1 GPA — Final Cum SAT-V .58 .50 (N = 318) SAT-M .43 .46 In a sample of 520 graduating seniors only, coefficients for Final Cum were .26 for SAT-V and .22 for SAT-M; yr 1 coefficients were not shown. The cumulative GPA distribution in the graduating sample was markedly restricted (SD = .46) relative to the cum GPA distribution for the entire entering cohort (SD = .75).

Table 1 (cont.). Studies Involving the Prediction of Cumulative GPA for Periods Exceeding the Freshman Year, with or without Comparative Coefficients for Shorter-term GPA and/or Other Criteria (e.g., major-field GPA, graduation/nongraduation)

Source	*Variables*[a]	*Salient Findings*[b]
Wilson (1976) Samples of graduating seniors, six liberal art colleges; selectivity indicated by range of SAT-V from (approximately) mean = 650, SD = 65 through mean = 575, SD = 90	Predicted Yr 1 GPA Yr 1 GPA Cumulative GPA 4-yr (five colleges) Senior-yr GPA (one college)	Predicted Yr 1 GPA vs. College A* B C* D* G H* Yr 1 GPA .53 .44 .43 .58 .30 .49 4-yr Cum .48 .35 .35 – .28 .49 Senior yr .32 N 724 327 403 356 236 215 *These are averages (unweighted) of coefficients from two classes. *Note:* These are simple correlations between the GPA criteria and a weighted composite of SAT-V, SAT-M, HSR, and CB Achievement Tests; weights were developed in earlier samples.
Farver, Sedlacek, and Brooks (1975) Two cohorts of black students entering Univ of Maryland and samples of white students from same cohorts; analysis by race and sex	Predictive composites of SAT-V, SAT-M, and HSR: composites weighted by data for cohort A were used to obtain composite-criterion correlations in cohort B, and vice versa.	Yr 1 GPA / 2-yr Cum / 3-yr Cum / No. of cases Black M .504 .392 .276 (102, 102, 79) Black F .610 .599 .610 (137, 110, 77) White M .615 .555 .540 (149, 111, 89) White F .612 .635 .575 (151, 125, 102) *Note:* For each subgroup, these are weighted averages of cross-validated coefficients for two cohorts. Note decrease in N due to attrition.
Price and Kim (1976) Junior and senior business students at Fort Hays Kansas State College; random sample	HSR ACT Age Cumulative GPA (2- or 3-yr)	Predictor / Cum GPA Beta weights HSR .267 ACT .552 Age .189 N = 93 R = .84 (no cross-validation)
Lunneborg (1977) University of Washington, cohort entering in 1971, members graduating in four years	HSR Washington Precollege Tests English Usage Quantitative Skills Others 4-yr Cum GPA Graduation/Nongraduation	Predictors / 4-yr Cum GPA / Grad/Nongrad HSR .51 .23 English Usage .42 .14 Quantitative Skills .34 n.a. HSR + "best others" (R) .57 .25 (N = 1,600+ survivors of an entering cohort of about 3,000 students)
Hardesty (1980) Cohorts entering DePauw University in 1973, 1974, or 1975, GPA records accumulated through fall 1979	HSR SAT-V SAT-M Final cumulative GPA (based on from one to four years)	Final Cum GPA Correlation by major areas Predictor / Final Cum / Sci / Human / Soc Sci / Educ / Nursing HSR .57 .59 .58 .60 .52 .41 SAT-V .43 .39 .49 .50 .38 .46 SAT-M .40 .43 .41 .43 .47 .47 Multiple (R) .625 .66 .65 .67 .60 .62 N (1,758) (260) (333) (545) (144) (88)

Table 1 (cont.). Studies Involving the Prediction of Cumulative GPA for Periods Exceeding the Freshman Year, with or without Comparative Coefficients for Shorter-term GPA and/or Other Criteria (e.g., major-field GPA, graduation/nongraduation)

Source	*Variables*[a]	*Salient Findings*[b]
Wilson (1978, 1980) State University B; cohorts of minority and nonminority students entering in 1971, records as of spring, 1978; SAT-V mean = 492, SD = 101 for nonminority students and 411 and 110 for minority students	HSR SAT-V, SAT-M Sem 1 GPA Yr 1 GPA Final Cum GPA Progress–highest level attained within six years: 7 = Postgraduate enrollment 6 = BA/BS on schedule 5 = BA/BS behind schedule 4 = Senior status 3 = Junior status 2 = Sophomore status 1 = Freshman status	Total sample analysis* Predictor / Sem 1 GPA / Yr 1 GPA / Final Cum GPA / Progress SAT-V .40 .44 .48 .22 SAT-M .41 .44 .46 .26 HSR .51 .55 .56 .34 PrMinSI** .55 .59 .61 .35 PrMinCu** .55 .59 .61 .34 Sem 1 GPA – .88 .80 .48 Yr 1 GPA – – .86 .48 Ns per coefficient ranged from 1,173 to 1,235 due to missing observations; Ns for analyses involving Sem 1, Final Cum, and Progress were essentially constant. *Correlational gradients for minority and nonminority samples were similar; level of validity was higher in the minority than in the nonminority sample. **Regression-weighted composites of V, M, and HSR, based on minority-sample data, for predicting Sem 1 GPA (PrMinSI) and final Cumulative GPA (PrMinCu). When nonminority equations were used correlations were essentially identical.
Wilson (1978, 1981) Selective liberal arts college; cohorts of male minority and nonminority students entering in 1971, 1972, 1973, and 1974; records as of spring 1978. Graduates who completed no work elsewhere. Nonminority SAT-V mean = 644, SD 72; minority SAT-V mean = 528, SD 88 For the sample under consideration, graduation rates were approximately 89 percent for nonminority students and 92 percent for minority students.	SAT-V, SAT-M HSR CB Achievement Average Yr 1 GPA 2-yr Cum GPA 3-yr Cum GPA 4-yr Cum GPA "Longest stay" Cum GPA (last Cum GPA of record–Yr 1, 2-yr, 3-yr or 4-yr Cum)	Total sample analysis* Predictor / Yr 1 GPA / 2-yr Cum / 3-yr Cum / 4-yr Cum / Longest-stay Cum / N/Coefficient** SAT-V .422 .442 .468 .495 .460 (964, 955, 928, 701, 964) SAT-M .375 .391 .401 .400 .387 (964, 955, 928, 701, 964) HSR .358 .361 .358 .336 .345 (906, 898, 873, 672, 906) CB Achievement .517 .528 .542 .536 .526 (941, 933, 906, 687, 941) Composite predictor Pred Yr 1. Min .569 .582 .586 .578 .573 Pred 3 yr cum.Min .561 .574 .576 .567 .563 (874, 867, 845, 655, 874) Pred Yr 1. Non .574 .590 .597 .593 .585 Pred 3 yr cum.Non .572 .589 .597 .596 .587 *Correlational gradients were similar in minority and nonminority samples; the level of correlation was lower in both subgroups than in the total sample, and lower in the minority than in the nonminority sample. **Cohort entering in 1974 was in second semester of senior year, hence no 4-yr cum GPA was available, thus accounting for low N for this criterion. #Four composite predictors based on regression of the designated criteria (Yr 1 or 3-yr cum GPA) on the four admissions variables in minority (Min) and nonminority (Non) samples.

[a]In this listing of variables, "HSR" refers to a measure of high school performance (e.g., rank in class or high school GPA). "CB" refers to College Board and "ACT" to American College Tests. "SAT" designates the CB Scholastic Aptitude Test-Verbal (V) and Mathematical (M). Unqualified references to "cumulative GPA" are from sources in which, by inference, the GPA was based on four years of study (e.g., the sample was comprised of graduates only). "Final cumulative GPA" is a variable-period GPA that may reflect from one semester to four years of study. This final undergraduate cumulative GPA is available for essentially all members of an entering freshman cohort, unlike fixed-period cumulative GPA indices (e.g., 2-yr cum, 3-yr cum, or 4-yr cum) that are available only for individuals who complete the number of years indicated. More generally, the list of variables is not in every instance an exhaustive enumeration for a given study. Only those variables that were judged to be most directly relevant for admissions-related consideration were listed.

[b]Unless otherwise noted, the correlation coefficients in the tables are zero-order coefficients. Relationships between independently computed GPA for a given period (semester or year) and a cumulative GPA reflect self-correlation.

ally identical predictive power and that the conclusion applied to the prediction of both two-year and four-year average grades.[4]

- Similar consistency may be seen in the results of the several smaller-scale studies in Table 1, regardless of the particular pattern of criterion selection and sample definition involved.

Several studies examined the relationship of a *predicted freshman average grade* to both the freshman and a cumulative GPA with generally consistent results, indicating that a composite that is best-weighted for predicting the freshman-year average also predicts the cumulative average at about the same level of accuracy (see Willingham, 1962; Hills, Bush, and Klock, 1964; Wilson, 1976; Farver, Sedlacek, and Brooks, 1975). Evidence of the essential interchangeability of predictive composites regardless of the particular general GPA criterion employed in their development is provided in studies by Wilson (1978, 1980, 1981).

- For example, at one anonymous state university, predicted Semester 1 GPA and predicted final cumulative GPA, based on data for a sample of minority students only, had essentially equal validity patterns with respect to both short-term and long-term general GPA and other criteria in the total sample under consideration, minority plus nonminority. A similar pattern occurred in analyses at a selective liberal arts college.

The foregoing lines of reasoning and the associated empirical evidence provide strong support for the working hypothesis that admissions variables are likely to be as valid for predicting longer-term cumulative GPA criteria as for predicting the usual first-year GPA criteria—that for purposes of assessing the predictive validity of admissions variables the two criteria are statistically interchangeable. In addition, it is evident that use of long-term cumulative GPA criteria introduces problems not involved in the use of the freshman-year GPA. First, the time interval between admission and the collection of criterion data would necessarily be longer. To obtain a final cumulative GPA, for example, a four-year interval would be involved during which there may be changes in institutional practices with respect to grading, promotion or loss of standing, course distribution requirements, and curricular patterns, all of which could affect the nature of the long-term cumulative GPA. Moreover, use of a fixed-base, four-year cumulative would involve validation in samples that are not representative of the entering cohort; for example, point biserial coefficients in the .4 range for cumulative GPA vs. graduation/nongraduation have been reported (French, 1957, 1958; Humphreys, Levy, and Taber, 1973). Variable-base final cumulative would not represent the same amount of work for all individuals in the entering cohort, etc.

These considerations lend support to the traditional practice of using freshman-year GPA rather than longer-term cumulative GPA in basic predictive validity studies.

[4]Schrader (1971) concluded later that these findings, plus those of the two previously cited studies, definitely supported the practice of using the first-year or two-year average as the criterion measure in basic validity studies, rather than the four-year average.

Table 2. Distributions of Shrunken Multiple Correlation Coefficients for Selected Elements of the Washington Pre-College Test Battery vs. Year 1 and 4-year Cumulative GPA, by Subject

Yr 1 GPA in 39 Subject Areas	*Four-yr cum GPA in 54 Subject Areas*
.7 55	.7 1
.6 00124 77789	.6 13456 79
.5 00001 11123 56688	.5 00112 22345 66789
.4 34456 789	.4 02234 45555 66778 99
.3 117	.3 11223 44678 9
.2 8	.2 268
N = 39	N = 54
Median R = .51	Median R = .47

Note: In these "stem-leaf" distributions, the first digit of the correlation coefficient designates a class interval (e.g., .7 represents coefficients between .70 and .79). The traditional "N per interval" is replaced by recording the second digit of each coefficient in each interval. Considering the Year 1 GPA distribution, for example, two coefficients are in the .7 interval, namely, .75 and .75; in the .6 interval we read coefficients of .60, .60, .61, .62, . . . , .69, for a total of 10 coefficients.

Minimum N per coefficient was 200. The coefficients were adjusted for shrinkage. The medians of .51 and .47, by subject area, may be compared with medians of .65 for both Year 1 and 4-year Cumulative GPA (general university GPA).

Source: Tabulated from Langen (1965), Table XVI.

Predicting Cumulative GPA in Subject Areas

Admissions-related validation research has focused, appropriately, on predicting general GPA criteria reflecting a student's average level of performance without regard to the particular pattern of courses taken. In validation research concerned with educational guidance rather than selective admission, however, the overall GPA criterion typically is abandoned in favor of a number of specialized or "unbundled" GPA criteria reflecting performance in particular areas of study only. Standard tests of admissions and other tests of developed cognitive abilities and the high school record are moderately correlated with more narrowly defined cumulative GPA criteria (e.g., junior-senior-year GPA in the major field; cumulative GPA in particular subjects or subject areas over up to four years) as indicated by results of several of the studies summarized in Table 1 (see University, 1956; French, 1957, 1958; Hardesty, 1980).

- Table 2 summarizes results of a study by Langen (1965) of the comparative validity of elements of the Washington Pre-College Test Battery for predicting GPA in a large number of specific subject areas cumulated for the first year and over four years of

study.[5] Among other things, Langen was concerned with determining whether any practical differences would result if a GPA based on only one or two years of study in a given subject were substituted for a GPA accumulated over four years. On the basis of the results summarized in Table 2, Langen concluded as follows: "The existing four-year extent of achievement for (subject) criterion areas does not provide more reliable prediction, as one might expect, than would be provided from using a one- or two-year extent of achievement. Since there is little difference between the reliability of predictions (of subject GPA over different time periods of accumulation), the shorter period might be preferred since it reduces maturation time for validation procedures" (pp. 173-74). It is of interest to note that the median multiple correlations for four-year and one-year subject GPA criteria (.47 and .51, respectively) were substantially lower than the reported values for the corresponding general GPA criteria at the University of Washington (R = .65 for both Year 1 and final cumulative).

- Exhibit A provides salient data from a study by Schoenfeldt and Brush (1975) of the intercorrelations of GPAs in each of 12 subject areas, accumulated over 13 quarters following matriculation at the University of Georgia, and distributions of simple validity coefficients with respect to these criteria for SAT-V, SAT-M, and the high school GPA. Correlations of these predictors with grades in all the subject areas were positive, but (as at the University of Washington) were lower than their validities for predicting the general GPA in units of the University of Georgia System (see Fincher, 1974; Hills, Bush, and Klock, 1964).

Predicting Independent Post-freshman-Year GPA

Available empirical evidence, although limited, provides rather consistent support for the proposition (a) that admissions measures tend to be less valid for predicting given independent post-freshman-year GPA (e.g., for the sophomore or senior year) than for predicting the freshman-year or cumulative GPA and, more generally, that there is a gradual, consistent, time-related decrease in predictive validity as GPA in successive years or terms (semesters or quarters) from the freshman through the senior year are introduced as criteria, and (b) that although the freshman-year GPA tends to be consistently more valid than admissions measures for predicting successive post-freshman-year GPA, the time-related decreasing validity gradient observed for admissions variables also occurs when post-freshman-year prediction is based on the freshman-year GPA (or other early GPA such as for the first quarter or semester).

Independent Studies

Results of three very early studies (Rogers, 1925; Anderson and Spencer, 1926; Freeman, 1931) and a more recent study (Juola, 1966) are summarized graphically in Figure 3, which shows trends in the relationship of freshman-level scores on several different cognitive ability measures to independently computed GPA in successive years, semesters, or quarters. Three of the studies also provided information regarding the relationship of an early GPA (freshman-year, first-quarter GPA) to the same later-independent GPA criteria). Although differing in detail, trends across all studies are similar.

- Rogers (1925) computed two sets of test/GPA correlations, one using data for all surviving students at the end of each successive year (a changing N or missing data sample) and one using only data for graduating seniors (a constant N or complete data sample), the latter to control for the effect of restriction of range due "to the elimination of poor students (which) tends to make the student body more homogenous and thus to decrease the amount of correspondence between test scores and academic marks." Trends in both samples were similar.

  Reflecting on the declining test/GPA correlational gradient, Rogers surmised that "industry and interest in academic subjects may count more towards success in the later years, or more mental adaptability may be necessary in the first year than in succeeding years," and concluded that interest was probably "... the most potent factor in decreasing the amount of correspondence between test scores and GPA between the freshman and the senior year." Although (in the missing data sample) the correlation of Year 1 GPA with GPA in Years two, three, and four declined, Rogers reported adjacent year correlations of .75, .74, and .72 for Years 1-2, 2-3, and 3-4, respectively, and an average of all interyear correlations of .68.
- In another early study involving test/GPA correlations over four college years, based on a missing data sample, Anderson and Spencer (1926) reported a declining validity gradient. They noted the same tendency, though to a lesser extent, in samples limited to those who remained four years. By way of explanation, they cited both (a) a regular decrease in the variability of academic grades over four years (under both complete and missing data conditions) *and* (b) the decreasing correlational gradient between freshman grades and grades of later years (shown in Figure 3 for a constant N sample), a trend reflecting instability in GPA standing across successive college years. They concluded that this latter trend "... would indicate

[5]For information about the Washington Pre-College Testing Program and related studies see, for example, Horst (1959); Dvorak and Langen, 1962; Lunneborg, 1966.

Exhibit A. Intercorrelations, Means, and Standard Deviations of GPAs in 12 Curricular Areas, High School GPA, and SATs, and Distributions of Correlations Between the Admissions Variables and the Curricular GPA

Intercorrelations, Means, and Standard Deviations of GPAs in 12 Curricular Areas, High School GPA, and SATs[a]

Variable	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	$\bar{X}$	*SD*	*Median Number Courses*[b]
1. Fine Arts	965	.38	.47	.41	.36	.44	.39	.32	.40	.35	.29	.27	.34	.25	.17	2.94	.75	1.95
2. Language	573	949	.50	.43	.49	.49	.19	.35	.32	.36	.29	.27	.44	.22	.21	2.49	.91	2.84
3. Humanities	961	942	1868	.59	.52	.69	.44	.50	.44	.53	.47	.34	.47	.43	.26	2.53	.71	6.08
4. Biological Sciences	634	657	1081	1085	.51	.63	.56	.48	.37	.45	.34	.34	.41	.30	.28	2.38	.86	2.46
5. Physical Sciences	890	866	1695	1029	1710	.55	.45	.46	.41	.37	.32	.33	.43	.18	.31	2.37	.91	3.59
6. Social Sciences	948	933	1814	1073	1682	1834	.39	.59	.47	.58	.44	.39	.45	.40	.28	2.55	.77	6.00
7. Agriculture	39	33	129	108	126	123	133	.22	.66	.31	.44	.38	.38	.19	.17	3.08	.80	2.77
8. Business	164	177	563	228	539	560	60	566	.31	.23	.35	.28	.36	.23	.23	2.62	.71	7.05
9. Education	443	281	640	403	600	637	17	144	643	.46	.36	.21	.37	.15	.14	3.16	.69	2.49
10. Home Economics	186	96	233	154	226	233	11	33	142	233	.36	.27	.34	.34	.20	3.05	.64	1.99
11. Speech/ Journalism	537	409	1055	545	987	1049	86	497	477	165	1059	.22	.33	.23	.13	2.88	.68	1.62
12. Physical Educ/ROTC	963	940	1852	1074	1694	1817	132	560	637	233	1053	1876	.31	.06	.09	2.92	.54	5.94
13. High School Grade Point	950	931	1825	1060	1671	1792	126	546	631	233	1036	1837	1861	.20	.17	3.11	.54	
14. SAT-Verbal	949	928	1818	1055	1664	1785	125	540	628	233	1032	1831	1853	1854	.41	505	83	
15. SAT-Math	950	929	1819	1056	1665	1786	125	540	628	233	1032	1832	1854	1854	1855	531	78	

[a]N for each mean and standard deviation appears on the diagonal. N for each correlation appears below the diagonal.
[b]Median number of courses being averaged, based on students taking one or more courses in the area.

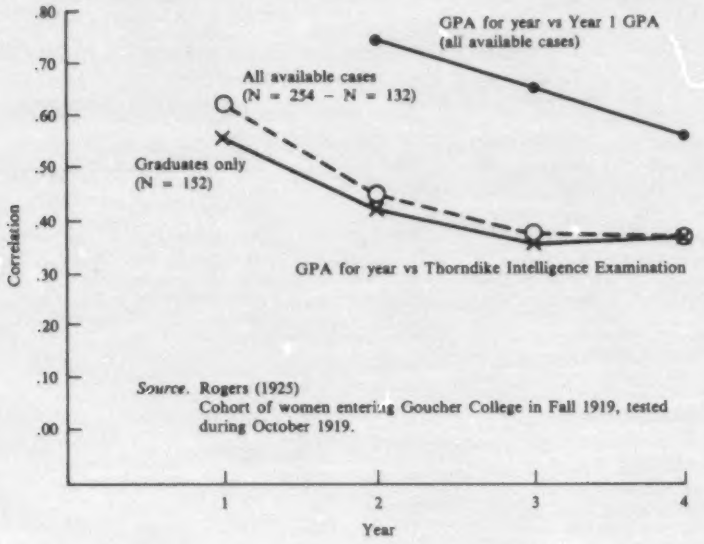
Source: Schoenfeldt and Brush (1975)
80 percent sample, 1968 freshmen, University of Georgia

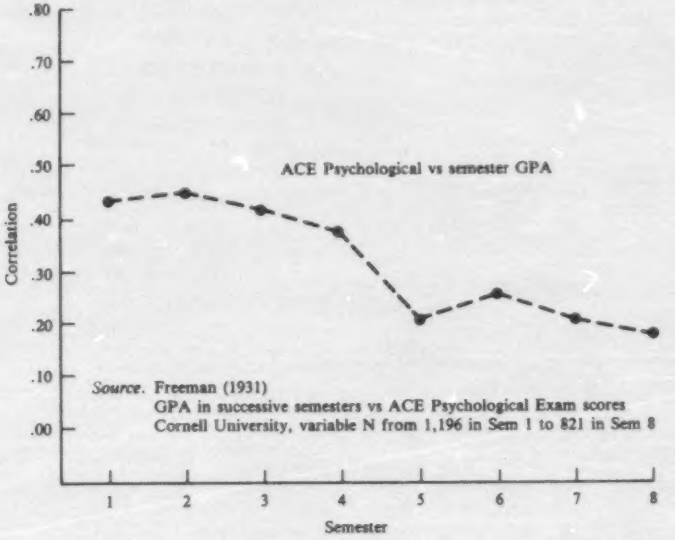
Distributions of Simple Correlation Coefficients for HSR, SAT-V and SAT-M, and GPA in All Courses Taken in 12 Subject Areas During 13 Quarters Following Matriculation

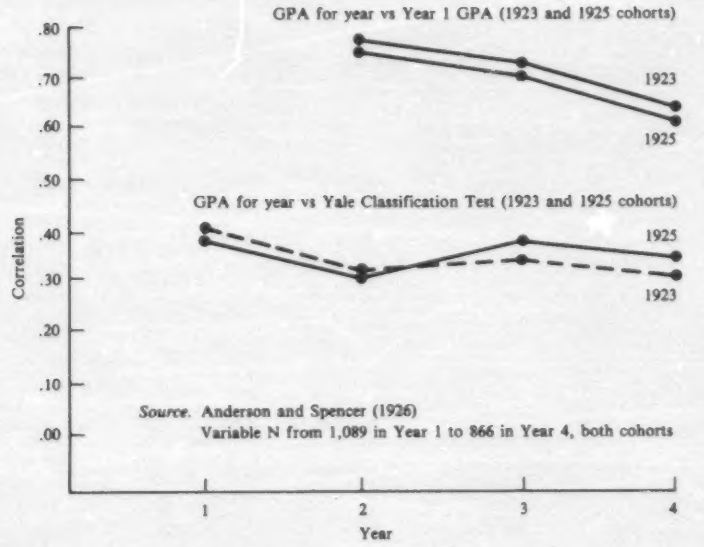
HSR	*SAT-V*	*SAT-M*
.6	.6	.6
.5	.5	.5
.4 13457	.4 03	.4
.3 13446 78	.3 04	.3 1
.2	.2 2335	.2 01368 8
.1	.1 589	.1 3477
.0	.0 6	.0 9
Median .375	.235	.205

Source: Schoenfeldt and Brush (1975)

Figure 3. Trends in the relationship of entry-level test scores to independently computed GPA in sucessive years, semesters, or quarters, and of earlier to later undergraduate GPA


GPA for year vs Year 1 GPA
(all available cases)
All available cases
(N = 254 – N = 132)
Graduates only
(N = 152)
GPA for year vs Thorndike Intelligence Examination
Correlation
Year
Source. Rogers (1925)
Cohort of women entering Goucher College in Fall 1919, tested during October 1919.


ACE Psychological vs semester GPA
Correlation
Semester
Source. Freeman (1931)
GPA in successive semesters vs ACE Psychological Exam scores
Cornell University, variable N from 1,196 in Sem 1 to 821 in Sem 8


GPA for year vs Year 1 GPA (1923 and 1925 cohorts)
1923
1925
GPA for year vs Yale Classification Test (1923 and 1925 cohorts)
1925
1923
Correlation
Year
Source. Anderson and Spencer (1926)
Variable N from 1,089 in Year 1 to 866 in Year 4, both cohorts

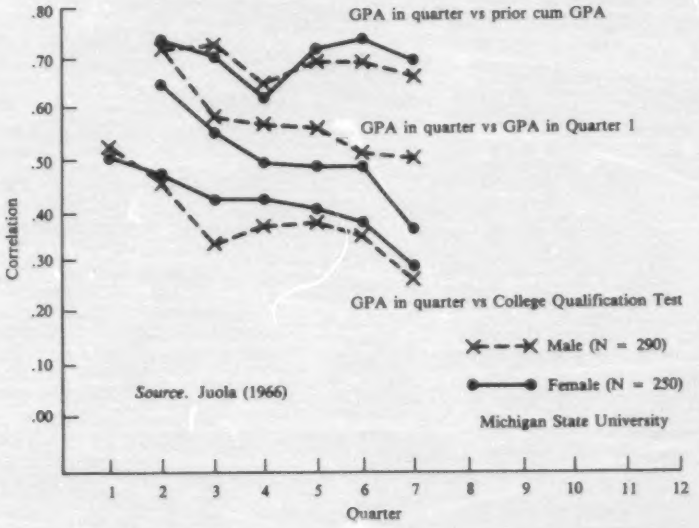

GPA in quarter vs prior cum GPA
GPA in quarter vs GPA in Quarter 1
GPA in quarter vs College Qualification Test
Male (N = 290)
Female (N = 250)
Michigan State University
Correlation
Quarter
Source. Juola (1966)

Exhibit B. Correlational Data Exhibiting a Declining Gradient in Relationships Between Repeated Measures in a Series as the Time Interval Between Measures Increases

Table 1. Intercorrelation of Mental Ages of Boys at Various Chronological Ages

(First Test)

	8	9	10	11	12	13	14	15	16	17	18
8		*721*	712	747	729	657	598	648	652	556	*444*
9	721		751	721	714	696	634	615	609	588	509
10	712	751		816	769	704	726	738	699	604	543
11	747	721	816		859	787	745	810	802	736	638
12	729	714	769	859		854	778	786	806	775	732
13	657	696	704	787	854		864	785	770	780	754
14	598	634	726	745	778	864		839	778	750	765
15	648	615	738	810	786	785	839		868	778	744
16	652	609	699	802	806	770	778	868		848	788
17	556	588	604	736	775	780	750	778	848		*828*
18	444	509	543	638	732	754	765	744	788	828	

Table 2. Intercorrelations of Trials of Complex Coordination Task

(from Fleishman and Hempel, N = 197)

	1	2	3	4	5	6	7	8
1		*75*	73	66	64	57	69	*59*
2	75		85	85	84	79	77	79
3	73	85		85	83	79	81	79
4	66	85	85		90	88	86	85
5	64	84	83	90		90	87	86
6	57	79	79	88	90		85	86
7	69	77	81	86	87	85		*90*
8	59	79	79	85	86	86	90	

Table 5. Intercorrelations of Weight of 275 Girls at Various Chronological Ages

	7	8	9	10	11	12	13	14	15	16
7		890	880	835	810	793	755	773	744	732
8	890		920	896	871	858	825	812	771	759
9	880	920		932	906	882	840	818	773	756
10	835	895	932		958	936	892	842	777	755
11	810	871	906	958		967	921	866	790	762
12	793	856	882	936	967		954	892	816	775
13	755	825	840	892	921	954		944	880	839
14	773	812	818	842	866	892	944		953	916
15	744	771	773	777	790	816	880	953		965
16	732	759	756	755	762	775	839	916	965	

Source: Humphreys (1967)

Table 11. Intercorrelations of Four Hour Exams on High School Mathematics

(Data from M. Beberman, N = 600)

	1	2	3	4
1		72	70	69
2	72		75	72
3	70	75		76
4	69	72	76	

Table 12. Intercorrelations of Grades in Four Semesters of Beginning Foreign Language

(N = 613)

	1	2	3	4
1		71	61	60
2	71		70	65
3	61	70		75
4	60	65	75	

Table 13. Intercorrelations of Grades in the Eight Semesters in Engineering

(N = 269)

	1	2	3	4	5	6	7	8
1		54	42	37	40	35	31	33
2	54		56	50	47	42	32	34
3	42	56		58	60	48	43	48
4	37	50	58		53	55	31	30
5	40	47	60	53		66	56	46
6	35	42	48	55	66		56	50
7	31	32	43	31	56	56		55
8	33	34	48	30	46	50	55	

Table 14. Intercorrelations of Semester Grades in the Liberal Arts College

(N = 314)

	1	2	3	4	5	6	7	8
1		59	56	53	47	37	42	43
2	59		52	51	48	40	41	45
3	56	52		53	51	47	35	41
4	53	51	53		51	56	47	48
5	47	48	51	51		57	48	50
6	37	40	47	56	57		47	46
7	42	41	35	47	48	47		63
8	43	45	41	48	50	46	63	

that the college grade is a variable function in some degree independent of any stable trait measured either by the classification test or by the marking system of the freshman year" (p. 306).

- Freeman (1931) followed the same line of inquiry with respect to test validity using independently computed semester GPA in a missing data sample at Cornell University, and found a similar declining validity gradient.
- Juola (1966) found the declining validity gradients portrayed graphically in Figure 3 for both test scores and first-quarter GPA with respect to GPA in seven successive quarters in a constant N sample at Michigan State University, noting that these trends were similar to those he had observed earlier in a missing data sample.[6]

These early studies provided evidence of a declining gradient in the predictive validity of both entry-level test scores and the freshman-year (or other early) GPA as GPA in successive years or terms are introduced as criteria, and that, in the case of test scores, the decline could not be accounted for by progressive restriction of range on the predictor due to the selective elimination of low-scoring students. Decreasing GPA variability across successive years was noted as a factor having the potential for lowering correlations of predictors with grades in later college years. None of the studies used the high school record as a predictor and none employed a currently used admissions test (although all the tests involved measured general developed cognitive abilities). No general conceptual framework linked the varied studies.

Studies by Humphreys and Associates within the Framework of the Simplex

The most thorough and consistent effort to elaborate and explicate the declining validity phenomenon is represented by the work of Humphreys (1960, 1967, 1968, 1976a, 1976b); Humphreys and Taber (1973); and Lin and Humphreys (1977). Humphreys (1960) provided a conceptual framework for the declining validity phenomenon by calling attention to the fact that matrices of intercorrelations of undergraduate semester GPA, like matrices for repeated measures in other time series under conditions involving growth, maturation, or learning (e.g., repeated measures of cognitive ability, performance on successive shorthand tests during a given learning period), conformed generally to the simplex pattern—that is, at all points in a simplex matrix, the correlation between earlier and later measurements decreases regularly as the time interval between measurements increases. In a later paper, Humphreys (1967) provided a number of matrices conforming generally to the simplex pattern: for example, intercorrelations of repeated measures of mental age at annual intervals between ages 8 and 18, eight trials of a complex coordination task, grades in four semesters of beginning language study, and GPA in eight semesters in engineering and liberal arts samples, as shown in Exhibit B.

Table 3. Correlation Between Repeated Measures in a Series: Initial vs. Final Measure, First vs. Second Measure, and Last Two Measures in the Series

		Measures Being Correlated		
Variable or Measure	*Series*	*1st vs. Last in Series*	*1st vs. 2nd in Series*	*Last Pair in Series*
Mental ages of boys at chronological ages	Ages 8–18	.44	.72	.83
Trials of complex coordination task	Ages 1–8	.59	.75	.90
Standing height of girls	Ages 7–16	.81	.99	.99
Weight of girls	Ages 7–16	.73	.89	.96
Vocabulary scores	Grades 2–6	.56	.65	.76
Reading scores	Grades 2–6	.58	.61	.73
Spelling	Grades 3–6	.56	.52	.69
Semester grades in engineering	Sem 1–8	.33	.54	.55

Note: Data are adapted from Humphreys (1967).

- These matrices conform only generally to the simplex pattern because the gradient of decrease in correlation between earlier and later measures is not completely consistent—some reversals in the gradient occur. However, the general pattern of lower correlation with increases in the time-interval between measures is evident, and each matrix is characterized by higher correlation between the first and second pair than between the first and last pair of measures in the series under consideration. Table 3 highlights these tendencies by showing only correlations between the first and last, the first and second, and the last pair of measurements in each of several series, *all with constant N*, from Humphreys (1967).

In evaluating the decreasing correlational gradient for the semester GPA data, Humphreys observed that "... on the basis of logic alone [the] predictive validities of tests given at the beginning of the period [i.e., college] will decrease in a similar fashion," and called attention to Juola (1966), whose findings provided empirical confirmation.

In a landmark study extending this line of inquiry into the prediction of college GPA, Humphreys (1968) related scores on basic predictors (high school rank and the ACTP

[6] Juola extended the line of inquiry involved in these studies by asking whether or not freshman-level test scores contribute anything to predictions made on the basis of earlier grades alone. He found that scores on the College Qualification Test did not make an incremental contribution to the prediction of GPA in later quarters after taking into account GPA in earlier quarters. His findings with regard to this particular line of inquiry will be considered in greater detail, along with similar findings from other contexts, in a subsequent section.

battery) to GPA in each of eight semesters in a large sample of freshman entering the University of Illinois in 1962 and 1963, first using a missing data (changing N) matrix and then using a complete data (constant N) matrix including only data for graduating seniors. In the correlational matrices studied, each separate correlation was a weighted average of the elements of correlation matrices computed separately by sex and by the college in which each student was enrolled, thus controlling for possible differences by sex and college. The following findings are noteworthy:

- In both the constant and the changing N samples, the intercorrelations of semester GPA generally followed the simplex pattern *and both ACT scores and high school rank exhibited a declining gradient from the first through the eighth semester.* The constant N sample controlled experimentally for restriction of talent, so the results could not be explained by the attrition among low ability students.
- Using adjacent semester correlations as lower bound estimates of reliability of the semester GPA, Humphreys concluded that "the near constancy of the adjacent semester correlations rules out substantial differences in reliability of freshman and senior grades as an explanation of markedly reduced predictor correlations."[7]
- The standard deviations of semester GPA, although decreasing from the first through the eighth semester in the missing data or changing N samples, were essentially the same in each semester for the sample of graduating seniors.

In explaining these data, Humphreys (1968) leaned toward the hypothesis that people are changing and that "aptitude for college work" is far from stable, as opposed to the hypothesis that the GPA criterion is changing. He noted that it would be desirable to have ACT retest data at the end of the senior year in order to determine whether senior-level scores would correlate more highly with senior-level GPA than with earlier GPA—a necessary condition for the "people are changing" hypothesis (Humphreys, 1960).

Later, Humphreys and Taber (1973), concluding that obtaining retest information on ACT from graduating seniors was not feasible, determined postdictive validities[8] for the GRE Aptitude Test and GRE Advanced (achievement) tests in a student's major, using data for a large number of University of Illinois students (including transfers) who normally take these tests during the senior year. The following findings were noteworthy:

- In a constant N sample, both GRE verbal and GRE quantitative scores exhibited *postdictively* a higher correlation with freshman than with senior grades; coefficients for the Advanced Tests were somewhat lower for the last four than for the first four semesters. The postdictive correlational gradients for these senior-level tests were generally similar to the predictive validity gradients for ACT and the high school record previously reported.
- In an Advanced Test analysis, employing a composite matrix of within-group correlations in which the elements were weighted averages of elements of matrices computed for groups classified by major field of the Advanced Test taken, the verbal and quantitative measures postdictively exhibited somewhat higher correlations with earlier than with later semester GPA, but the coefficients for the Advanced Tests followed a different gradient: they exhibited a slightly rising gradient from Semester 1 through Semester 4 and a slightly declining gradient from Semester 5 through Semester 8, and the correlation for Semester 8 GPA was lower than that for Semester 1 GPA. Figure 4 shows the predictive validity gradient for ACT in the Humphreys (1968) sample, and postdictive correlations for GRE-Verbal and GRE-Advanced Tests (from Humphreys and Taber, 1973), in differing samples from the University of Illinois.

Humphreys and Taber (1973) concluded from these findings that "variations of the explanation that the criterion is changing should certainly be considered" as opposed to a "people are changing hypothesis," in any effort to explain the observed similarity in predictive and postdictive validity gradients, citing somewhat similar findings of a more limited study by Lunneborg and Lunneborg (1970).

- Lunneborg and Lunneborg (1970) obtained high-school-level/college-senior-level test-retest data on a battery of standard ability tests for a small sample of senior volunteers (N = 126) at the University of Washington. The pattern of correlation with GPA in each of 12 successive quarters was similar for both the high-school-level and college-senior-level test scores. Trends in the Lunneborg and Lunneborg (1970) data are illustrated in Figure 5. The post-

[7]The near constancy of adjacent semester correlations led Humphreys to believe that the reliabilities across semesters were equal. In a later analysis of the Humphreys (1968) matrix, Werts, Linn, and Jöreskog (1978) used a procedure developed by Jöreskog to test whether the 1968 matrix formed a simplex and, if so, obtain reliability estimates (within the framework of the simplex model as suggested by Humphreys, 1960). They concluded that the correlations of eight semesters of GPA had a simplex form and that reliabilities across semesters were, in fact, equal. Evidence regarding GPA reliabilities across years or semesters is quite limited. Etaugh, Etaugh, and Hurd (1972), using different methodology, reported lower reliabilities for senior-year than for freshman-year grades. Questions regarding GPA reliabilities require local, empirical answers, of course.

[8]These are coefficients involving circumstances in which the performance of individuals on an independent variable is determined *after*, rather than before, their performance on a dependent or criterion variable.

Figure 4. Gradients in correlation with GPA by semester of tests administered near the beginning and near the end of college

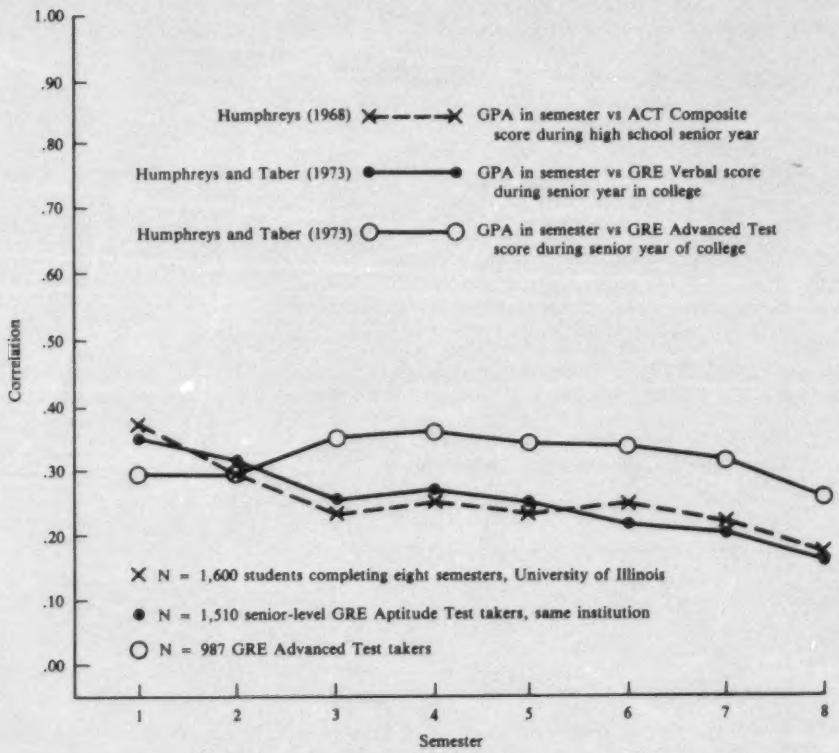
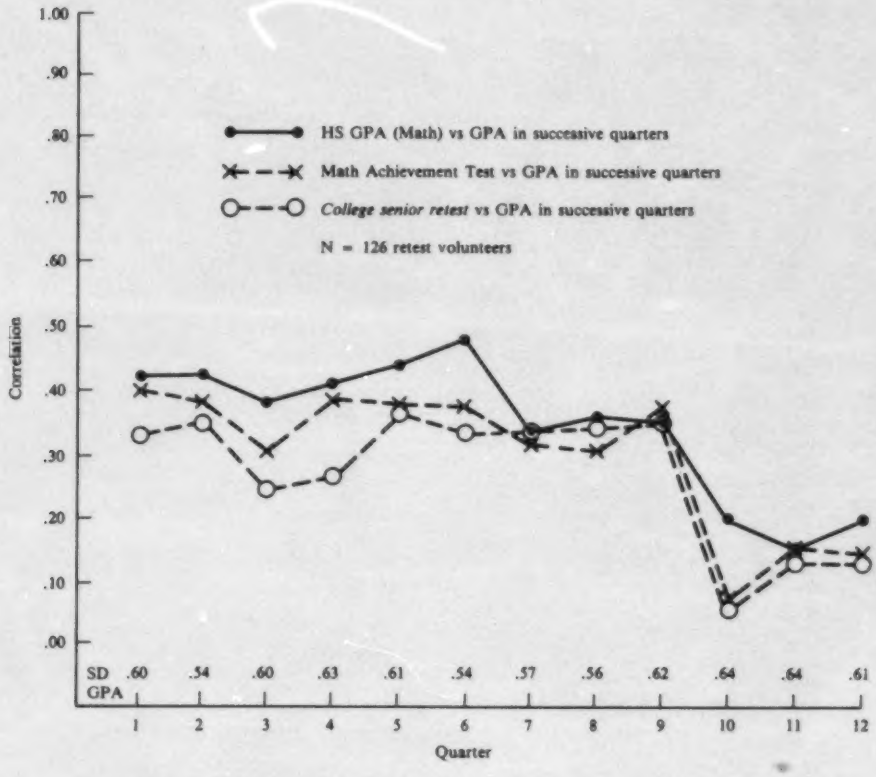


Figure 5. Predictive and postdictive correlations with GPA in successive quarters



Source. Lunneborg and Lunneborg (1970)

dictive correlations of the senior-level test (mathematics achievement) with successive quarter GPA paralleled the predictive validity of the high-school-level scores and the predictive validity of the high school record. The correlational gradients, per se, were quite irregular overall and were characterized by a sharp drop in both predictive and postdictive correlations with GPA in the last three quarters. This result was associated with sharply diminished interquarter correlations during the senior year (for quarters 10, 11 and 11, 12, correlations were .37 and .48, respectively, compared to average adjacent quarter correlations of approximately .6 for earlier marking periods).

In a later study in the simplex series, Lin and Humphreys (1977) analyzed correlations between undergraduate and graduate GPAs as well as between these GPAs and standard graduate and professional school admissions tests in large samples (N = 1,200) of law school and graduate students (in mathematics, physics, and chemistry).

- Among other things, they found (a) that the undergraduate GPA matrices, computed by year rather than by semester for both samples, conformed to the simplex pattern, and (b) that in the graduate school sample, GRE Advanced as well as Aptitude scores, correlated more highly with freshman grades than with senior grades, unlike the irregular pattern observed by Humphreys and Taber (1973), previously plotted in Figure 4.[9]

These investigators concluded that available evidence required consideration of both "changes in people" and "changes in the GPA criterion" in any general explanatory rationale for the simplex phenomenon, per se, and the associated declining validity of admissions measures.

Ancillary Evidence Bearing on the "People Are Changing" Hypothesis. Evidence regarding the degree of stability of performance on tests of general cognitive tests during and prior to the college years is pertinent to an assessment of the "people are changing" vs. the "criterion is changing" hypothesis. A considerable amount of stability may be inferred from observed findings of comparable predictive and postdictive validity gradients for measures such as the ACT battery and the GRE Aptitude and Advanced Tests. Humphreys and Taber (1973) did not report ACT/GRE correlations. Lunneborg and Lunneborg (1970) reported a stability coefficient, corrected for attenuation due to unreliability, of .91 for the mathematics achievement test for which high-school-level predictive and college-senior-level postdictive validity gradients were shown in Figure 5. Other evidence regarding the stability of cognitive measures over various periods of time during the college year is provided in Table 4.

Test-retest coefficients have been reported for ACT scores over time periods ranging from two weeks through four years (see Table 4). ACTP (1973), for example, reported stability coefficients ranging between .86 and .89 for intervals up to 20 weeks and two-year stabilities of .84. Dumont and Troelstrup (1981) reported a four-year correlation of .91 for ACT Total, using parallel forms test-retest data in a sample of seniors in one Tennessee university.

Findings by Humphreys and Taber (1973) and Lin and Humphreys (1977) of comparable predictive and postdictive validity gradients for ACT and GRE tests suggest not only a relatively high degree of stability over four years for the types of abilities measured by both tests, but also a relatively high degree of overlap between the two different cognitive measures. There appears to be no evidence regarding the concurrent or the long-range relationship between ACT and GRE Aptitude or Advanced Test scores. However, in unpublished analyses from a study of postsecondary outcomes for the national high school (National Longitudinal Study) Class of 1972 (Creech et al., 1977), it was found that the multiple correlation of SAT verbal and mathematical scores with ACT Total was approximately R = .85, among academic curriculum students in the NLS sample.

Schrader (1967) reported a correlation of .85 between SAT verbal and GRE verbal scores and of .84 between SAT mathematical and GRE quantitative scores in a sample of college-senior-level GRE candidates identified as having taken the SAT as high school seniors by matching files maintained by ETS. For liberal arts students, the GRE verbal score mean was about 20 points higher than expected from SAT verbal mean, and for engineering students, the GRE quantitative mean was about 30 points higher than expected from the SAT mathematical mean.

Other studies of SAT/GRE relationships over a four-year period have reported somewhat lower coefficients than those found by Schrader, but they were based on much more highly selected samples.

- Nichols (1964) reported correlations of .66 and .76 for SAT-V/GRE-V and SAT-M/GRE-Q in a small, highly selected sample of National Merit Scholars (N = 356); the SAT-V and SAT-M standard deviations were 52 and 70, compared to standard deviations of 89 and 98 for SAT scores in the Schrader study. Nichols reported curricular effects similar to those reported by Schrader—students who majored in engineering, chemistry, and mathematics exhibited increase in Q relative to V, while those who majored in English and arts and humanities increased in V relative to Q.
- Wilson (1970) found (a) correlations of .76 and .78 for SAT-V/GRE-V and SAT-M/GRE-Q in a larger (N = 878) and slightly less highly selected sample (seniors in several women's colleges), with SAT standard deviations of 73 and 80, and (b) curricular

[9]Because of the highly selected nature of the law school sample, the postdictive validity gradient for the Law School Admission Test with undergraduate GPA was characterized by coefficients that were generally either very low or slightly negative.

Table 4. Correlational Data on the Stability of Standing on Admissions Tests Over Various Time Intervals

Admissions Measures	N	Time Interval Between Measures[a] 2-6 Weeks	12 Weeks	20 Weeks	1 Year	2 Years	4 Years	Source of Data
ACT Composite								*ACT (1973)*
Study 1	433	.86						Scholarship nominees, reg. admin-retest
Study 2	1,659	.88						Scholarship nominees, reg. admin-retest
Study 3	1,268	.89						Scholarship nominees, reg. admin-retest
Study 4	3,089		.89					ACT repeaters, regular admins.
Study 5	2,428			.87				ACT repeaters, regular admins.
Study 6	63					.84		Test-retest, 2-year college
Study 7	972					.84		Regular admin-retest, five colleges
Study 8	112						.91	*Dumont and Troelstrup (1981)* Regular admin-retest, state university
Scholastic Aptitude Test								*Alderman (1981)*
Study 9 SAT-V	1,325				.88			SAT repeaters, regular admin.
SAT-M					.87			
SAT/GRE Aptitude								*Schrader (1967)*
Study 10 SAT-V/GRE-V	468						.85	National adminis., matched records,
SAT-M/GRE-Q							.84	college seniors, 15 colleges
Study 11 SAT-V/GRE-V	265						.65	*Nichols (1964)*[b]
SAT-M/GRE-Q							.76	Merit Scholars study, national admins.
Study 12 SAT-V/GRE-V	878						.76	*Wilson (1970)*[b]
SAT-M/GRE-Q							.78	Seniors, eight selective colleges national admins.
American Psychological Exam								*McConnell (1934)*
Study 13	71						.83	Test-retest, liberal arts college

[a]For a sample of 217 GRE candidates identified as having repeated the GRE Aptitude Test four times in regular national administrations, correlations between the first and the fourth administrations (with varying intervening time intervals) were reported as .823 for GRE-Verbal and .810 for GRE-Quantitative (Rock and Werts, 1979).

[b]None of the coefficients in this table have been adjusted for differences between sample and population standard deviations. These samples are more highly selected than the Schrader sample. For example, standard deviations for SAT-V and SAT-M were 52 and 70 for the Nichols sample and 73 and 80 for the Wilson sample, as compared to 89 and 98 for the sample studied by Schrader. Using the standard deviations reported by Schrader (which in turn are smaller than general SAT candidate standard deviations) as population values, coefficients corrected for restriction become .826 and .881 (SAT-V vs. GRE-V) and .914 and .893 (SAT-M vs. GRE-Q) for the Nichols and Wilson data, respectively.

effects similar to those reported by Schrader and Nichols.

- Using the SAT standard deviations reported by Schrader as "population" values (these were smaller than general SAT candidate standard deviations), coefficients corrected for restriction on the SAT become .826 and .881 (SAT-V vs. GRE-V) and .914 and .893 (SAT-M vs. GRE-Q) for the Nichols and Wilson data.

No four-year stability data appear to be available for the SAT itself. However, one-year stability coefficients of .88 and .87 (SAT-V and SAT-M) were reported by Alderman (1981) for candidates identified as repeaters in national SAT administrations. It seems probable that four-year stability coefficients for the SAT would be at least as great as those reported for the SAT vs. the GRE or for ACT.

Generally speaking, these findings suggest collectively a high degree of overlap among different admissions measures as well as considerable stability in the types of abilities being measured by such tests over the 17 to 22 year age range in samples of college students. This ancillary evidence on the stability of general cognitive abilities, along with evidence of comparable predictive and postdictive validity gradients for such measures, lends support to the hypothesis of gradual change in the GPA criterion as at least a partial explanation of the simplex pattern of GPA correlations and the declining validity gradients observed.

Additional evidence bearing on the changing people vs. the changing criterion hypothesis is provided in studies showing rising gradients in predictive validity with respect to a fixed, remote college GPA criterion (Semester 1 GPA) of (a) successive GPA in Grades 1 through 8 (Lewis, 1966), and (b) scores on a statewide achievement test battery administered at intervals in Grades 4 through 12 (Loyd, Forsyth, and Hoover, 1981). Salient characteristics and findings of these two studies are summarized in Exhibit C.

- Lewis (1966) obtained Semester 1 GPA and elementary school GPA data for a small, changing N sample composed of graduates of a single suburban high school whose elementary schooling had been completed in the same district. Twenty-four colleges provided Semester 1 GPA data. Correlation of GPA in Grades 1 through 3 with college Semester 1 GPA was negligible, but GPA correlations tended

Exhibit C. Studies involving the prediction of college GPA from GPA in elementary grades and from very early test performance

Predicting College GPA from GPA in Elementary School and Early Test Performance

Lewis (1966)

24 colleges and universities reported to a middle class suburban high school the 1st semester grades of three successive graduating classes. Data for individuals whose elementary schooling had been completed in the same district, grades 1 through 8, were analyzed.

Variables: GPA in Grades 1–8; High School GPA; Henmon Nelson Test of Mental Ability, (administered in Grade 8, or at about age 14)

Correlation of predictors with first-semester college GPA

Predictor	r	N
Grade 1 GPA	.05	(60)
Grade 2 GPA	.11	(61)
Grade 3 GPA	.00	(64)
Grade 4 GPA	.30	(68)
Grade 5 GPA	.28	(75)
Grade 6 GPA	.33	(78)
Grade 7 GPA	.41	(81)
Grade 8 GPA	.52	(84)
Henmon score	.36	(84)
H. S. GPA	.49	(85)

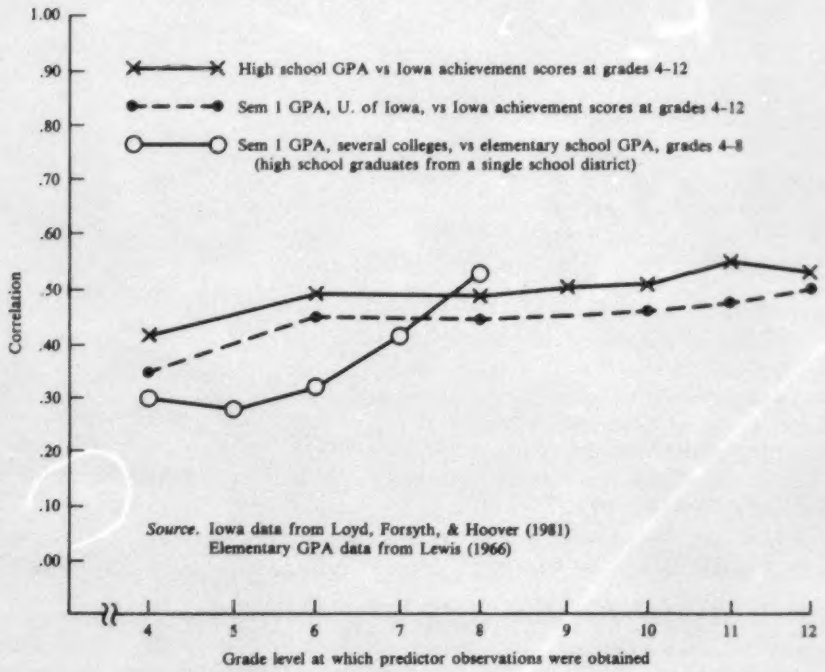
Loyd, Forsyth, and Hoover (1981)

Approximately 2,000 students entered the University of Iowa in 1977. Study based on those with scores on high school achievement tests administered in grades 9, 10 11, and 12 and elementary 4, 6, and 8 in Iowa schools.

Variables: High School GPA; Iowa Tests of Basic Skills (elementary); Iowa Tests of Educational Development (high school level); ACT Composite; (Also subtests); Sem 1 GPA

Correlation of composite scores with first-semester GPA and HSGPA

Grade level composite	HSGPA (n)	r	Sem 1 GPA r	(n)
4	(1012)	.42	.35	(976)
6	(1246)	.49	.45	(1209)
8	(1215)	.49	.44	(1181)
9	(752)	.50	.45	(728)
10	(931)	.51	.46	(899)
11	(1053)	.55	.48	(1019)
12	(653)	.53	.50	(631)
ACT Composite		—	.46	(1912)



Source. Iowa data from Loyd, Forsyth, & Hoover (1981)
Elementary GPA data from Lewis (1966)

to increase steadily from Grades 4 through 8, from r = .30 to r = .52; the Grade 8 GPA correlated somewhat better with Semester 1 GPA than did the high school average or an ability test administered in Grade 8, at estimated age 14.

- Loyd, Forsyth, and Hoover (1981) obtained data for a sample of students entering the University of Iowa in 1977, which included relatively large numbers of individuals who had taken the Iowa Tests of Basic Skills in Grades 4 through 8 and/or the Iowa Tests of Educational Development in Grades 9 through 12. Both high school GPA and Semester 1 GPA were available for substantial numbers of these students. A generally rising predictive validity gradient may be seen in Exhibit C for the achievement tests over Grades 4 through 12, with respect to both high school and Semester 1 GPA.

It is evident that over the age-grade ranges involved in these two studies, predictor-criterion correlations tended to exhibit a rising gradient as the time interval between testing and criterion measurement decreased. Over the same age-grade range, the correlation between repeated general cognitive measurements has been shown to follow the simplex pattern, as indicated by the intercorrelations of mental age (MA) between ages 8 and 18, from Humphreys (1967), previously summarized in Exhibit A and Table 3.

- It may be seen that the correlation between MA at ages 8 and 18 was .44, but for ages 17 and 18 the correlation was .83.
- It is reasonable to hypothesize that the simplex pattern of MA intercorrelations reflects changes in the rank ordering of people between ages 10 and 18 due to differential patterns of growth in cognitive skills. If so, it may be inferred that "changes in people" underlie the increase in predictive validity for successive general achievement measures with respect to a fixed, remote criterion—as the "trait" being measured by the tests became increasingly more stable between ages 8 and 18, the correlation of the measure with the criterion increased.[10]

It is noteworthy that corrected stability coefficients approaching .9 have been found for admissions measures over the college age range in samples of college students. In any event, the ancillary evidence examined in this section suggests that the decreasing correlation between admissions tests and GPA over the college years is not due primarily to major changes in the rank-ordering of students in basic academic abilities and, thus, lends support to the "criterion is changing" hypothesis.

[10]The line of reasoning involved here is based in part on an assumption that intercorrelations of scores on the general achievement batteries involved would conform to the same general simplex pattern as that for mental age.

Other Validity Studies

Further evidence regarding patterns of correlations of standard admissions measures with GPA in successive semesters is provided in data from two quite different undergraduate settings, namely, an anonymous state university (Wilson, 1978, 1980) and a highly selective liberal arts college for men (Wilson, 1978, 1981), in studies not specifically concerned with the declining validity phenomenon, per se, but rather with comparative assessment of the long-term performance in college of minority and nonminority students. In these studies, the primary target for prediction was the cumulative GPA (see the summary of validities for predictors and composite predictors with respect to cumulative GPA criteria in Table 1). However, matrices of intercorrelations of predictor and semester GPA were generated for both "constant N" samples (graduating students only) and "changing N" samples.

Table 5 summarizes selected findings from these two settings for samples of graduating seniors only (all students, minority and nonminority);[11] coefficients from Humphreys (1968), based on data for graduating seniors, are included for comparative purposes. In evaluating trends in correlation, it is important to note that unlike Humphrey's correlation matrices, the state university correlation matrices were not controlled by sex and school of enrollment (at this particular university, students elect majors in one of nine different schools after completing freshman-level work in a general university division), nor was the selective college matrix controlled for students' major fields. There were also more missing GPA data in the latter two samples than in the Illinois sample.

The following points are noteworthy:

- In all three samples, for the high school record and test predictors alike, correlations with GPA were lower over the last four semesters; in both the Illinois and the liberal arts samples, there was a gradual decrease in correlation from the first through the eighth semester, with only modest reversals.
- In the anonymous university sample, however, the correlation of both SAT verbal and SAT mathematical tended to increase with successive GPA through the fifth semester, and to decrease steadily from the fifth through the eighth semesters.
- Standard deviations of GPA variables in the university sample were essentially constant across semesters—0.59 (Semester 1), 0.61, 0.64, 0.63, 0.65, 0.64, 0.64, 0.62 (Semester 8). In the selective college sample, standard deviations fluctuated somewhat more—1.37 (Year 1), 1.65 (Semester 3), 1.44, 1.51, 1.49, 1.48, 1.38 (Semester 8)—but no marked decrease in GPA dispersion was evident.

[11]Trends in the sample including all students (nongraduates and graduates) were similar.

Table 5. Data Indicating a Declining Gradient in the Correlation of Various Admissions Measures with Independently Computed GPA in Successive Undergraduate Semesters in Restricted Samples Including Only Seniors or Graduates

Setting	*Predictor*	*Independent Semester GPA Criterion* 1	2	3	4	5	6	7	8	*Study*
University (Illinois) (N = 1600)[a]	HS Rank	39	34	28	27	24	26	24	22	Humphreys (1968)
	ACT Composite	38	30	24	26	24	25	22	17	
Anonymous university (N = 530)[a]	HS Rank	51	45	45	41	45	40	37	36	Wilson (1978, 1980)[b]
	SAT-Verbal	42	42	46	48	47	40	38	37	
	SAT-Mathematical	42	40	47	45	49	37	30	32	
Selective liberal arts college (N = 950)[a]	HS Rank	36[c]	–	29	28	27	25	19	20	Wilson (1978, 1981)
	CB Achievement Av.	51[c]	–	45	40	42	39	40	34	
	SAT-Verbal	42[c]	–	40	34	39	35	35	33	
	SAT-Mathematical	38[c]	–	36	29	32	28	28	20	

Note: Data tabled are correlation coefficients (without decimal).

[a]Although samples include only individuals who successfully persisted through the senior year, some semester GPA values are missing. These are representative Ns.

[b]Coefficients tabled are from analyses not included in the referenced reports.

[c]Correlation of predictor with Year 1 GPA; separate Sem 1 and Sem 2 GPA not available.

Table 6. Correlations Between GPA in Adjacent Undergraduate Semesters in Restricted Samples Including Only College Seniors or Graduates

Type of Setting	*N*	*Semester GPAs Correlated* 1, 2	2, 3	3, 4	4, 5	5, 6	6, 7	7, 8	*Study*
State Univ., engineering	91	69	65	65	62	61	65	72	Humphreys (1960)
University of Illinois	1,600[a]	56	49	56	51	55	54	54	Humphreys (1968)
State Univ., GRE takers	2,600[a]	64	62	63	61	65	61	60	Humphreys and Taber (1973)
SUNY College	79	69	54	63	62	74	56	N.A.	Schell (1975)
Anonymous university	530[a]	58	61	67	64	66	62	56	Wilson (1978, 1980)[b]
Selective liberal arts	950[a]	N.A.[c]	N.A.[c]	67	63	67	63	61	Wilson (1978, 1981)
Weighted mean		609	576	621	589	627	609	583	

Note: Tabled values are correlation coefficients without decimals.

[a]These samples include only individuals who persisted successfully through the senior year. However in some cases GPA values were not available for every semester. The Ns of 2,600, 530, and 950 represent roughly median Ns per coefficient; the N of 1,600 was more nearly constant. These Ns are used herein as though constant for the purpose of computing the weighted means, and they reflect general differences in sample size across the various studies.

[b]Only coefficients for the total sample are reported in the references cited.

[c]Separate Sem 1 and Sem 2 GPA were not available; only the Year 1 GPA was reported for this sample.

- The observed lower correlations for later semesters did not appear to be due to major changes in GPA reliabilities since adjacent semester GPA correlations were relatively stable. Table 6 shows adjacent semester GPAs for the anonymous university and selective college samples along with similar data from Humphreys (1968), Humphreys and Taber (1973), and two additional samples that were not employed in prediction studies.[12]

[12]Very few matrices of intercorrelations of semester GPA were located in the review. Adjacent semester correlations averaging about .6 for the limited data in Table 6 may prove to be typical. Rogers (1925) estimated adjacent year GPA correlations in the neighborhood of .7.

In both the anonymous university and the liberal arts settings, trends for comparatively small samples of minority graduating seniors (not tabled) were somewhat different from those reported in Table 5 for the total sample of graduating seniors (Wilson, 1978).

- In a small group of minority students completing eight consecutive semesters (N = 66) at the anonymous university, the correlations of a composite predictor, with GPA by semester, did not exhibit a declining gradient, being, respectively, .53, .54, .57, .54, .64, .58, .59, and .54 for semesters 1 through 8; for similar sample of nonminority students (N = 322), comparable coefficients were .56, .42, .53, .50, .49, .37, .40, and .39 (a generally

declining gradient). Unpublished data indicate that the standard deviations of GPA in the minority sample were considerably *higher* in the later than in the earlier semesters, while those for the nonminority sample were generally similar across semesters.

- In the selective college sample (without separate semester GPA for the freshman year), coefficients of a composite predictor for Year 1 GPA and successive semester GPA thereafter, in a sample of graduating minority students (N = 82 with some missing GPA), were as follows: .39 (Year 1), .27, .32, .35, .34, .33, and .11 (Semester 8 GPA); in the similar nonminority sample (N = 792, with some missing GPA), coefficients were .54 (Year 1), .50, .43, .46, .42, .41, and .35 (Semester 8). For the minority sample the validity of the composite predictor was highest for Year 1 GPA; thereafter the validity gradient was flat or slightly increasing from Semesters 3 through 7, but then dropped sharply for Semester 8 GPA, unlike the generally declining gradient for nonminority students. In the minority sample, there was also rather marked instability in GPA between the seventh and eighth semesters (r = .41 for Semester 7, 8 as compared, for example, to the minority Semester 6, 7 correlation of r = .56, and r = .61 for Semester 7, 8 GPA in the nonminority sample).

Only one other study located in this review reported a nondeclining validity gradient (like that for minority students in the anonymous university sample) for admissions variables vs. successive, independently computed post-freshman-year GPA. Siegelman (1971) studied the relationship of SAT-V, SAT-M, and the high school record to independently computed GPA in successive years for a small sample of men (N = 80) and women (N = 95) who had completed four years of study including a sequence of courses required for teacher licensing. SAT findings for men were anomalous (near zero). However, for both sexes, the validity of the high school average, and for women, the validity of the SAT scores, did not decline when independently computed GPA in successive years were used as criteria, and correlations with the final cumulative average were higher than with the first-year average.

The anonymous university minority student findings and those reported by Siegelman, both from quite small samples with many pertinent variables uncontrolled, represent minor exceptions to the general pattern of findings indicating that the validity of admissions measures (standard tests and the high school GPA) tends to decline as GPA in successive post-freshman-year study periods are introduced as criteria.

- It is instructive to note that in the small, highly selected sample of minority students, for example, the increase in GPA standard deviations across semesters was matched by a steady increase in the correlation between adjacent-semester GPA (from .48 for Semester 1, 2 to .74 for Semester 7, 8). Moreover, contrary to the simplex model, the correlational gradient between Year 1 GPA and GPA in successive subsequent semesters was essentially flat (coefficients of .45, .45, .45, .51, .44, and .44). And a best-weighted composite of the three admissions variables for predicting Semester 1 GPA in the *total* minority sample correlated .55 with Semester 1 GPA, .64 with Year 1 GPA, and .72 with the final cumulative GPA.
- In Siegelman's data, the correlation of admissions measures with independently computed GPA did not decrease from the freshman through the senior year.

In the present context, these findings are significant primarily because they reinforce an observation by Humphreys (1976a) that in a matrix of GPA data in which the elements do not conform to the simplex pattern, the correlations of predictors with *cumulative* averages should tend to increase gradually as the number of components increase rather than stabilize or exhibit a slightly rising gradient.

Predicting Later Grades from Earlier Grades

Available evidence indicates that the freshman-year GPA tends to be more valid than admissions measures (tests and the high school record) for predicting independently computed post-freshman-year GPA, but that the correlation of freshman-year GPA with GPA in successive years or terms tends to show a generally declining gradient similar to that for tests or for the high school record (see, for example, Figure 3). Further evidence regarding the relationship of Year 1 GPA and Semester 1 GPA to GPA in successive subsequent undergraduate semesters in several samples of graduating seniors is provided in Table 7. The data are from Humphreys (1960, 1968), Humphreys and Taber (1973), Schell (1975), and Wilson (1978, 1980, 1981).[13]

- These data indicate that Year 1 GPA is consistently more highly correlated than Semester 1 GPA with GPA in successive semesters, undoubtedly as a function of greater reliability. Validity coefficients for admissions variables, previously shown in Table 5 for the Humphreys (1968) and Wilson samples only, tended to have a similar gradient but at a consistently lower level that that for Year 1 GPA.
- The data in Table 7 also indicate that considerable change in rank ordering in terms of GPA takes place among "successful students" between the freshman

[13] For the Humphreys, Humphreys and Taber, and Schell references, the Year 1 coefficients tabled were calculated by the writer from the published correlation matrices. They represent the relationship of an equally weighted composite of Semester 1 and Semester 2 GPA to the designated semester GPA criteria. Schell (1975) reported a matrix of semester GPA correlations in a study concerned primarily with whether average grades awarded by the faculty in later semesters were higher than those in earlier semesters.

Table 7. Declining Gradients in Validity of Year 1 and Semester 1 GPA as Predictors of GPA in Successive Subsequent Undergraduate Semesters in Samples Including Seniors or Graduates Only

Type of Setting	*Sample Predictor*	*Semester GPA Criterion* 2	3	4	5	6	7	8	*Study*
State Univ., engineering	Sem 1 GPA	69	55	46	45	41	34	33	Humphreys (1960)
	Year 1 GPA[a]	–	65	57	54	55	41	42	N = 91
University of Illinois	Sem 1 GPA	56	46	44	40	42	39	34	Humphreys (1968)
	Year 1 GPA[a]	–	54	50	46	45	42	39	N = 1600[b]
State Univ., GRE takers	Sem 1 GPA	64	51	53	48	42	41	35	Humphreys and Taber (1973)
	Year 1 GPA[a]	–	62	61	55	50	48	40	N = 2600[b]
SUNY College	Sem 1 GPA	69	46	39	42	33	26	N.A.	Schell (1975)
	Year 1 GPA[a]	–	54	48	50	41	38	N.A.	N = 79
Anonymous university	Sem 1 GPA	58	54	50	49	46	37	34	Wilson (1978, 1980)
	Year 1 GPA	–	64	58	57	53	43	41	N = 530[b]
Selective liberal arts	Sem 1 GPA	–	–	–	–	–	–	–	Wilson (1978, 1981)[c]
	Year 1 GPA	–	69	64	63	58	57	46	N = 950[b]
Weighted mean	Sem 1 GPA	609	497	494	449	423	396	345	
	Year 1 GPA		610	580	539	501	471	408	

Note: Tabled values are correlation coefficients without decimals.

[a]This is an equally weighted composite of Sem 1 and Sem 2 GPA. The coefficients reported here were calculated from available Sem GPA data from the referenced studies in which Year 1 GPA was not treated as a variable.

[b]These samples include only individuals who persisted successfully through the senior year. However, in some cases semester GPA values were not available for all semesters. The Ns of 2,600, 530, and 950 are roughly median Ns per coefficient; the N of 1,600 was more nearly constant although some values were missing. These Ns were used as though constant for the purpose of computing the weighted mean, and they reflect general differences in sample size.

[c]Separate Sem 1 and Sem 2 GPA not available for this sample; only the Year 1 GPA was used.

and the senior years; accuracy of predicting a specific post-freshman-year GPA criterion from the freshman-year GPA decreases steadily over the undergraduate years.

However, based on the work of Willingham (1962) and Juola (1966), and special reanalyses of correlation matrices from Humphreys (1968) and Wilson (1978, 1980, 1981), accuracy of predicting both independently computed GPA in successive terms and cumulative GPA beyond a given term remains relatively stable across the undergraduate years *when prediction is based on the previously earned cumulative grade record.*

Using a matrix of intercorrelations among quarterly GPA averages for 390 students who completed 12 quarters at the Georgia Institute of Technology, Willingham (1962), in an extensive unpublished institutional analysis, estimated correlations between the cumulative GPA at the end of successive quarters and (a) the cumulative GPA over all remaining quarters (b) the senior-year GPA, and (c) the final cumulative GPA. Exhibit D shows the intercorrelation matrix[14] and plots of trends in the correlations, including the correlation of GPA in the first quarter with GPA in each subsequent quarter.

- The "o" graph in the exhibit reflects the accuracy of predicting all subsequent performance from all prior performance. The essentially flat correlation gradient indicates that accuracy of predicting student performance (GPA) through graduation from previous performance was about the same at all levels from the first quarter through the senior year.
- The "x" graph indicates that the correlation between cumulative GPA at the end of successive quarters and the final cumulative GPA gradually approaches unity as the earlier cumulative becomes an increasingly large element in, and ultimately becomes identical with, the final cumulative.

Juola (1966), as previously noted (see Figure 3 and related discussion), examined test/GPA and GPA/GPA relationships over the first seven quarters in a sample of 290 men and 250 women students randomly selected from the freshman cohort entering Michigan State University in fall 1960.

- For both men and women, Juola found that the best prediction of GPA in any given quarter was the cumulative GPA prior to that quarter; also, that accuracy of predicting GPA quarter by quarter from previous cumulative GPA did not decline markedly over the first seven quarters.

Both Willingham (1962, see also 1964) and Juola (1966) concluded that after taking into account the freshman-year or previous cumulative GPA, scores on tests

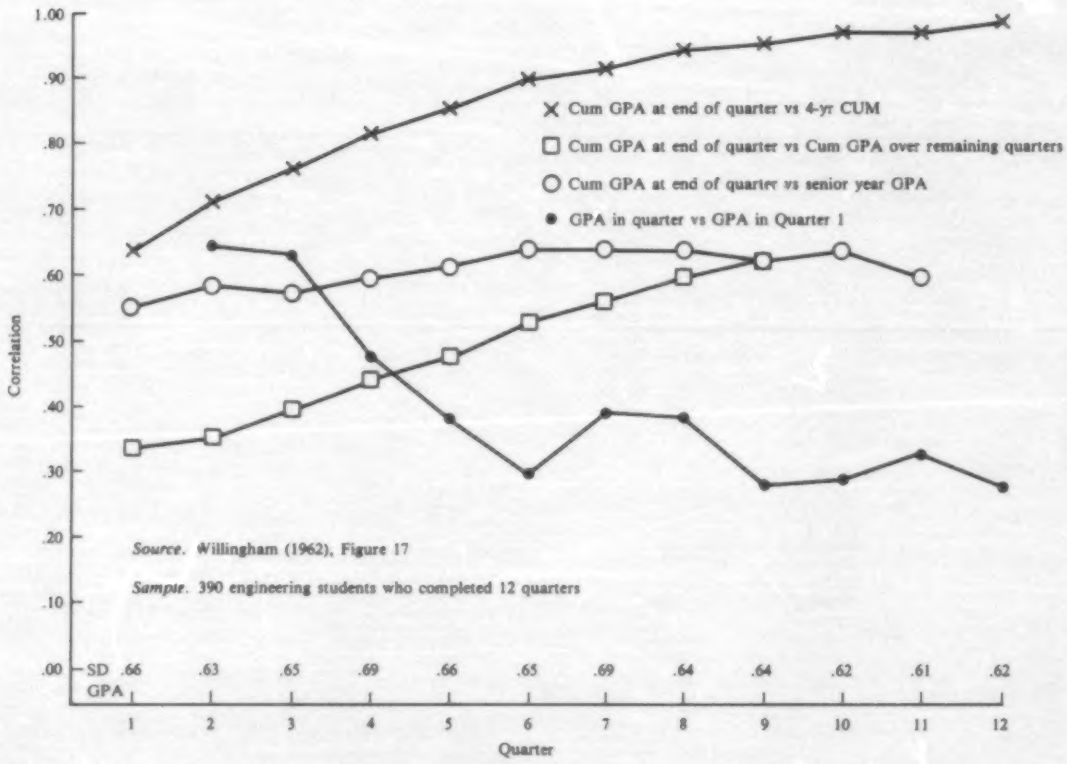
[14]The matrix included in Exhibit D represents the only correlational data located in this survey for GPA in 12 successive quarters. Note that the correlation between adjacent-quarter GPA was relatively unstable. Interquarter correlations were comparatively high in the first and in the last years and low during the sophomore and the junior years.

Exhibit D. The relationship of cumulative GPA at the end of successive quarters to later cumulative and senior-year GPA, and of first-quarter GPA to GPA in successive quarters, with the corresponding intercorrelation matrix for GPA in each of 12 successive quarters

Intercorrelations Among Quarterly Averages for 390 Students Who Completed Twelve Quarters.

Quarter	1	2	3	4	5	6	7	8	9	10	11	12
1												
2	.65											
3	.64	.62										
4	.46	.50	.49									
5	.35	.41	.41	.43								
6	.29	.36	.44	.47	.38							
7	.37	.38	.44	.44	.43	.47						
8	.39	.41	.45	.[illegible]	.39	.46	.42					
9	.28	.33	.3[illegible]	.39	.38	.39	.49	.50				
10	.29	.27	.30	.39	.38	.43	.45	.50	.44			
11	.33	.26	.37	.32	.38	.41	.44	.50	.49	.57		
12	.28	.30	.33	.31	.34	.39	.46	.44	.53	.59	.60	
Mean	2.4	2.3	2.2	2.2	2.2	2.3	2.3	2.4	2.5	2.5	2.5	2.6
SD	.66	.63	.65	.69	.66	.65	.69	.64	.64	.62	.61	.62

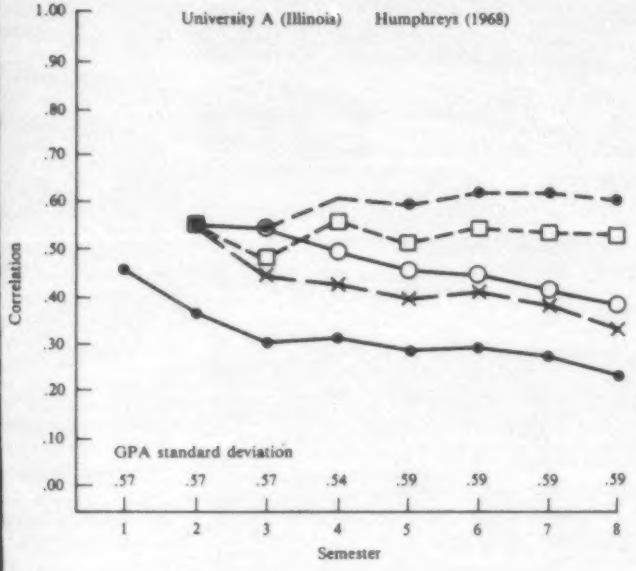
Source. Willingham (1962), Table 3A

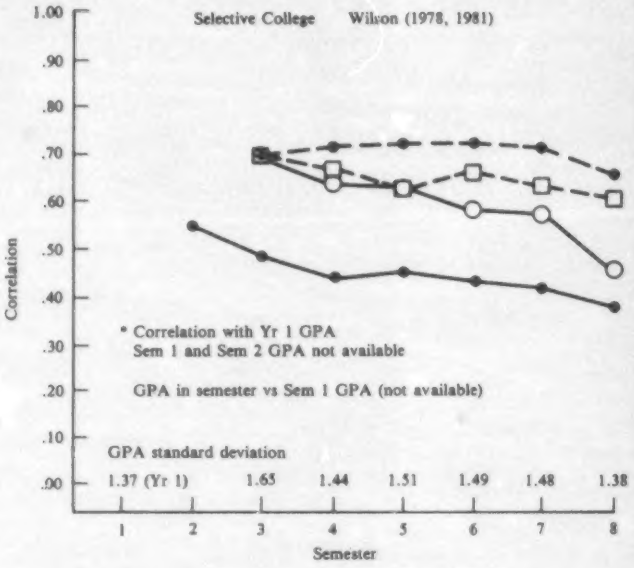

1.00
.90
.80
.70
.60
.50
.40
.30
.20
.10
.00
Correlation
X Cum GPA at end of quarter vs 4-yr CUM
□ Cum GPA at end of quarter vs Cum GPA over remaining quarters
○ Cum GPA at end of quarter vs senior year GPA
● GPA in quarter vs GPA in Quarter 1
SD GPA .66 .63 .65 .69 .66 .65 .69 .64 .64 .62 .61 .62
1 2 3 4 5 6 7 8 9 10 11 12
Quarter

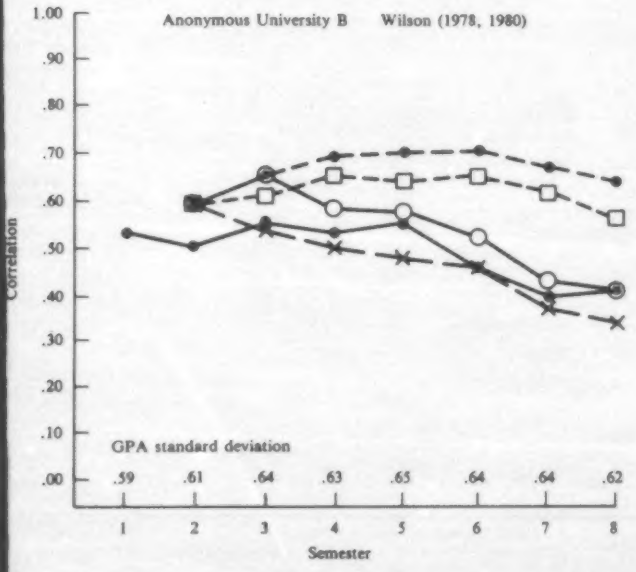
Source. Willingham (1962), Figure 17

Sample. 390 engineering students who completed 12 quarters

Figure 6. Trends in the correlation of GPA in successive undergraduate semesters with (a) all previous semester GPA, (b) GPA in the immediately preceding semester only, (c) Year 1 GPA, (d) Semester 1 GPA, and (e) a regression-weighted composite of available admissions measures, based on ad hoc supplementary analyses of correlation matrices, shown in Appendix A, for samples of seniors or graduates from the University of Illinois, anonymous university, and selective college.







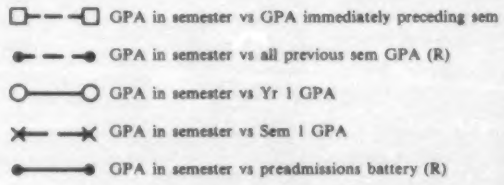

GPA in semester vs GPA immediately preceding sem
GPA in semester vs all previous sem GPA (R)
GPA in semester vs Yr 1 GPA
GPA in semester vs Sem 1 GPA
GPA in semester vs preadmissions battery (R)

Table 8. Studies Involving the Prediction of Independently Computed Post-freshman-year GPA Using both Earlier College GPA and Admissions Measures as Predictors

Study Context and Sample	*Variables*[a]	*Salient Findings*[b]
Miner (1927) Seniors, Arts and Sciences University of Kentucky (N = 102)	Army Alpha Sem 1 GPA GPA over last seven semesters (Sem 2–8 GPA)	*Predictor* Sem 1 Sem 2–8 N = 102 Alpha .50 .29 Sem 1 GPA – .58 Multiple R .58* *(Multiple computed from reported coefficients)
Weber, Brink, and Gilliland (1942) Graduate students at Northwestern U. whose undergrad study was also at Northwestern (N = 319)	Freshman-level centiles on ACE or Ohio State Psychological Cum UCPA GPA in graduate study (period not defined) Other variables	*Predictor* Cum UGPA Graduate GPA Freshman test .44 .31 UGPA cum – *.61** Test, UGPA (R) – .62 *(Note atypically high correlation between Undergrad GPA and Grad GPA in a sample with UG-to-Grad-school continuity)
Davis and Frederiksen (1954) Princeton University, liberal arts, students completing two years (N = 602)	SAT-V High School Record (HSR) First-semester GPA Yr 1 GPA Yr 2 GPA	*Predictor* Yr 2 *Predictor* Yr 1 Yr 2 Sem 1 GPA (S1) .74 HSR .62 .57 S-1, SAT-V (R) .74 SAT-V .40 .34 S-1, HSR (R) .76 HSR, SAT-V (R) .60
Spindt (1959) Third-year students at UCLA, both "native" and transfer; about one-third transfers (Total N = 833)	Lower division GPA Scores on an experimental College Board test battery including "aptitude" and "achievement" measures. GPA in Semester 5 (first semester of junior year). *Note:* Test battery was administered during junior year.	*Predictor* Lower division GPA vs. Sem 5 GPA, by department • Native students (10 departments) Lower-div GPA without tests Lower-div GPA with test range of coeff .51–.75 (r) .60–.77 (R)* average coeff .67 .698 • Transfer students (7 departments) range of coeff .53–.85 (r) .59–.90 (R)* average coeff .66 .70 *(Test coefficients not reported; multiples not corrected for shrinkage–no cross-validation).
Willingham (1962) Students completing at least three quarters at Georgia Institute of Technology (N = 799)	Predicted Grade (PGI), a composite of SAT-V, SAT-M, and HSR Yr 1 GPA Yr 2 GPA	*Predictor* Yr 1 Yr 2 PGI *.63* .43 Yr 1 *.63* Yr 2 PGI, Yr 1 (R) *(.631)*
Lewis (1964) State University of Iowa College of Pharmacy, students completing from one to four years	Iowa Testing Program Composite HSR*, Science Units, HSR(2)**; Yr 1 GPA, Yr 2 GPA, Yr 3 GPA, Yr 4 GPA *(GPA) **(Rank in class)	*Predictor* Preadm var* vs Yr 1 R = .602 (N = 107) Preadm var**, Yr 1 vs Yr 2 R = .608 (N = 87) Preadm var#, Yr 1, Yr 2 vs Yr 3 R = .652 (N = 70) Yr 3 GPA vs Yr 4 r = .705 (N = 63) *(HSR-2 and Test betas significant–simple r's not reported) **(Only test significant, but beta was negative [–.269]) #(None significant, *all betas were negative*)

Table 8 (cont.). Studies Involving the Prediction of Independently Computed Post-freshman-year GPA Using both Earlier College GPA and Admissions Measures as Predictors

Study Context and Sample	*Variables*[a]	*Salient Findings*[b]
Juola (1966) Michigan State University, students completing seven successive semesters (N = 540)	College Qualification Test (CQT) Quarter (Q) GPA Q1–Q7 Cumulative GPA at the end of successive quarters (Q1–Q2, Q1–Q3, . . . , Q1–Q6)	GPAs correlated* / r / Add Test (R) / CQT vs GPA** (r) / Q1 GPA vs GPA** (r) Q1 vs. Q2 / .705 / .713 / Q1 .52, .50 / –, – Q1–2 vs. Q3 / .712 / .713 / Q2 .46, .46 / .71, .64 Q1–3 vs. Q4 / .642 / .645 / Q3 .34, .42 / .58, .55 Q1–4 vs. Q5 / .694 / .697 / Q4 .38, .42 / .57, .45 Q1–5 vs. Q6 / .678 / .678 / Q5 .38, .40 / .56, .49 Q1–6 vs. Q7 / .659 / .664 / Q6 .35, .37 / .51, .49 Q7 .37, .28 / .50, .36 *Males only, N = 290; trends similar for females **first coeff for males, second for females (N = 350)
Leyman (1967) State University College at Cortland (NY), sophomore physical education majors	HSR Admissions Exam Year 1 GPA Year 2 GPA Scott Motor Ability Test	*Predictor* / GPA Yr 1 / GPA Yr 2 HSR (1) / .56 / .43 / Yr 1 (N = 201) Exam (2) / .41 / .32 / Yr 2 (N = 162) Yr 1 GPA (3) / – / .71 Multiple (1, 2)* / .60 / – Multiple (1, 2, 3)* / – / .71 *Includes Motor Ability Test
Humphreys (1968) University of Illinois, students completing eight semesters	ACT Composite HSR GPA Sem 1 through GPA Sem 8	
Wilson (1978, 1980) Anonymous state university, students from the September 1971 entering cohort who earned a bachelor's degree within six years	SAT-Verbal SAT-Mathematical HSR GPA Sem 1 through GPA Sem 8	Intercorrelation matrices from the Illinois State University and selective college samples were analyzed to determine the contribution of preadmissions variables and earlier postadmissions GPA variables to the prediction of (a) GPA in successive semesters and (b) GPA in Semester 8. Results of the analysis are presented separately.
Wilson (1978, 1981) Anonymous selective liberal arts college, male students from cohorts entering in 1970–1973 who graduated or reached senior status by spring 1977.	SAT-Verbal SAT-Mathematical HSR CB Achievement Test Average GPA Yr 1 GPA Sem 3 through GPA Sem 8	

[a]The listing of variables is not necessarily an exhaustive listing in every instance. "HSR" denotes an index of high school achievement (e.g., rank in class or high school GPA). "SAT" and "ACT" refer to the College Board Scholastic Aptitude Test and the American College Testing Program tests.

[b]The data presented provide a basis for assessing the contribution of preadmissions variables to the prediction of a post-freshman-year GPA criterion when an early postadmissions GPA, such as the Semester 1 or Year 1 GPA, is treated as an additional predictor.

and/or the high school record had no incremental validity for predicting subsequent GPA (a line of inquiry to be considered in detail later).

Supplementary Analyses

In order to extend the line of inquiry represented by the work of Willingham and Juola, special reanalyses of correlation matrices for samples of graduating seniors from Humphreys (1968) and Wilson (1978, 1980, 1981) were undertaken to determine the relationship to GPA in each of eight successive semesters of regression-weighted composites of (a) GPA for the preceding semesters and (b) the standard admissions variables employed.[15] The resulting multiple correlational gradients are plotted in Figure 6, which also includes plots of simple correlational gradients for first-year GPA, first-semester GPA, and GPA in the immediately preceding semester.

- While differing in detail, results for all three samples indicate that (a) the best prediction of GPA in any semester was provided by the prior cumulative GPA; (b) the second-best prediction was provided by the GPA in the immediately preceding semester and, consistent with findings of Willingham and Juola; (c) "continuous" GPA/GPA prediction of this nature is marked by considerable stability across semesters.
- In all samples, Year 1 GPA was a consistently better predictor of GPA in successive semesters than was the admissions battery. In the Illinois sample this was true for Semester 1 GPA as well. However, in the anonymous university sample, the admissions battery exhibited higher correlations with GPA in Semesters 3 through 8 than did the Semester 1 GPA. Generally speaking, the validity gradients for all these "fixed" predictors exhibited a declining gradient across semesters, although in the university sample, validities for the admissions battery did increase somewhat between the second and the fifth semester, decreasing thereafter.

Predicting Later GPA from Earlier GPA and Admissions Measures

Based on the evidence reviewed in the preceding sections, there is reason to believe that in general, the freshman GPA (earlier GPA) will predict later GPA with greater accuracy than test scores and the high school record considered separately and/or in best-weighted combinations. Very few investigators have used admissions measures and the freshman or other early GPA in studies of post-freshman-year prediction studies. Even fewer have been concerned directly with exploring the contribution of admissions measures to post-freshman-year prediction after taking into account a student's freshman GPA.

The evidence reviewed in this section indicates *that the prediction of post-freshman-year GPA is not improved by adding scores on commonly used admissions measures to the freshman-year (or other earlier) GPA.* Table 8 summarizes (in chronological order) the findings of (a) a small number of studies in which the investigator(s) directly assessed the contribution of standard tests, the high school record, or best-weighted composites, to the prediction of a post-freshman-year GPA criterion after taking into account the freshman or other early GPA, and (b) a study employing freshman-level tests and cumulative undergraduate GPA as predictors of graduate-level GPA for graduate students who completed all their undergraduate work at the institution of graduate enrollment. These findings are supplemented by the results of a special reanalysis (to be summarized later) of correlation matrices from the Illinois, anonymous university, and selective college settings that were generated in studies not concerned directly with the incremental validity question.

By far the most thorough examination of the relative contribution of admissions measures and the freshman or other early GPA to post-freshman-year prediction is provided in studies by Willingham (1962, 1964), and Juola (1966), previously cited.

- In his four-year longitudinal study involving a cohort of over 1,000 entering freshmen at the Georgia Institute of Technology, Willingham (1962) examined the incremental validity of preadmissions measures for predicting performance after the freshman year after taking the freshman GPA into account. He found that adding the Predicted Grade Index (a composite of SAT-V, SAT-M, and the high school record) to Year 1 GPA did not improve prediction of Year 2 GPA over that provided by the Year 1 GPA only.
- Willingham also examined the comparative utility of preadmissions measures and the freshman average for predicting each of three dichotomous post-freshman-year criteria, namely, final average (less than 2.0 vs. 2.0+), persistence (graduated or still enrolled in school vs. all others), and "success" (graduated or still enrolled with passing grades vs. all others). As may be seen in Exhibit E, the freshman average was the superior predictor with respect to each criterion.

Willingham (1962) concluded that the utility of the preadmissions measures appeared to be limited to a relatively short period of time before and after matriculation, even though they were moderately related to post-freshman-year performance. He later speculated that soon after admission, college grades may have tapped most of the variance in preadmissions measures that is useful for the purpose of

[15] The correlation matrices used for these special analyses and for additional supplementary analyses that are reported later are shown in the Appendix.

Exhibit E. The Relationship Between SAT Scores, High School Average (HSA), Predicted Grade (PG), and Freshman Average, and Various Measures of Performance.[a]

		SAT-V	*SAT-M*	*HSA*	*PG*	*Fresh. Average*
	N	1,059	1,059	1,059	1,059	932
Freshman Average	916	.40	.43	.48	.61	–
Sophomore Average	799	.27	.29	.36	.43	.63
Final Average (<2.0 vs. 2.0+)	1,046	.27	.36	.43	.50	.78
Persistance (Grad. and In School vs. Out of School)	1,059	.14	.24	.30	.32	.51
Success (Grad. or In Passing vs. All Others)	1,059	.19	.31	.35	.40	.64

[a]Each correlation is based upon the number of cases indicated in the row or column heading, whichever is smaller. Correlations based upon sophomore average are Pearson product-moment coefficients. All others are biserial coefficients.

Source: Willingham (1962), Table 9
Four-year longitudinal data for an entering cohort at the Georgia Institute of Technology

predicting post-freshman-year performance (Willingham, 1964).

Juola (1966), as noted earlier (see Figure 3 and related discussion), provided correlations between test scores and term grades over a seven-quarter span for a sample of 290 male and 250 female students selected randomly from the freshman class entering Michigan State University in fall 1960. He found that the cumulative GPA immediately preceding each term provided a considerably better basis for predicting GPA in that term than did either the *College Qualification Test* (CQT) or the first-quarter GPA. He then asked whether adding test scores to the cumulative GPA at the end of successive quarters would contribute to the prediction of GPA in the immediately following term (quarter).

- In the sample of males (shown in Table 8), the multiple correlations obtained when scores on the CQT were added to the cumulative GPA at the end of successive terms did not differ significantly (less than .001 on the average) from the simple correlations obtained for the cumulative GPA alone. Results for women were similar.

In a related line of inquiry, Juola found that in predicting grades in given courses from grades in previous courses *in the same area* (e.g., GPA in Communications 2 vs. GPA in Communications 1), test scores had only limited incremental validity. However, in predicting grades in a given course (e.g., Communications 1) from earlier grades in a course in a different area (e.g., Natural Science), test scores had considerable incremental validity. He concluded (a) that the prediction of later GPA generally was not enhanced by supplementing previous GPA with test scores when earlier and later GPA were comparable in content, but (b) that test scores might be useful supplements to previous GPA for predicting later GPA when there were content (curricular) differences involved–for example, in circumstances involving potential change of major field.

Tests and/or the high school record were also found to add nothing to prediction of various post-freshman-year GPA criteria in several more limited studies summarized in Table 7 (Miner, 1927; Davis and Frederiksen, 1954; Lewis, 1964; and Leyman, 1967). It is noteworthy that in one of these studies (Lewis, 1964), preadmissions measures were negatively weighted when included with earlier GPA in best-weighted composites for predicting Year 2 and Year 3 GPA. Such suppression is consistent with a "redundancy of information" thesis–all the later-GAP-related variance in the admissions measures was reflected in the earlier GPA.

Spindt (1959) examined the contribution of scores on a battery of experimental tests, administered during the junior year to samples of native and transfer students at UCLA, to the prediction of semester GPA after taking into account the cumulative lower-division GPA. The tests covered the fields of the social sciences, humanities, life sciences, physical sciences, and mathematics.

- In analyses involving native students in 10 departments, correlations of lower-division with Semester 5 GPA ranged from .51 to .75, with an average of .67; only one sample coefficient was below .61. Adding scores of the test battery yielded multiple correlations ranging from .60 to .77, with an average of .698. In seven departments, for samples of transfer students, results were similar–lower-division GPA/Semester 5 GPA coefficients ranged from .53 to .85, averaging .66, and adding test scores resulted in an improvement in the average to .70. Coefficients were not corrected for shrinkage. Spindt concluded that the overall improvement was too small to warrant further experimentation with the tests.

In a possibly unique study, Weber, Brink, and Gilliland (1942) examined the incremental validity of freshman-level test scores when added to the cumulative undergraduate GPA for predicting graduate GPA in a sample of Northwestern University graduate students who had completed all their undergraduate work at Northwestern. Analyses were conducted separately for students classified by the broad area of department of enrollment (e.g., natural sciences, social sciences, humanities) and for the total sample. Trends were similar in these analyses, and only the total sample results are shown in Table 8.

Table 9. Results of Special Reanalyses Indicating the Declining Incremental Contribution of Admissions Measures to the Prediction of a Remote Semester GPA Criterion as GPA in Successive Earlier Semesters Are Introduced as Predictors

	Correlation of Predictor(s) with Sem 8 GPA								
	University A (Illinois) (N = 1,600)			*Anonymous University B (N = 530)*			*Selective College (N = 950)*		
Predictor(s)	*Earlier GPA Only*	*Add Pre-Comp*	*Difference*	*Earlier GPA Only*	*Add Pre-Comp*	*Difference*	*Earlier GPA Only*	*Add Pre-Comp*	*Difference*
Preadmission composite (Pre-Comp)	–	241	–	–	406	–	–	384[a]	–
Sem 1 GPA	342	356	014	345	435	090	N.A.	N.A.	N.A.
Sem 1–2 GPA	386	392	006	416	466	050	463[b]	493	030
Sem 1–3 GPA	424	427	003	473	497	024	534	552	018
Sem 1–4 GPA	470	472[c]	002	510	524	014	560	575	015
Sem 1–5 GPA	519	521	002	573	578[d]	005	577	589	012
Sem 1–6 GPA	559	560	001	610[e]	613	003	611	621	010
Sem 1–7 GPA	612	613[f]	001	645[e]	648	003	660	667[g]	007

Note: The coefficients (decimal omitted) reported in this table were computed from intercorrelation matrices for samples comprised of individuals who completed eight semesters from University A (Humphreys, 1968), University B (Wilson, 1978, 1980), and Selective College (Wilson, 1978, 1981). The matrices were treated as "complete data" matrices, with Ns as reported herein, even though there was variation in N/coefficient due to the fact that some semester GPA data were missing. Preadmissions variables differ across studies: University A (ACT Composite, HS Rank), University B (SAT-V, SAT-M, HSR), Selective College (SAT-V, SAT-M, HSR, CB Achievement Average).

[a]SAT-M negatively weighted in this and subsequent composites.

[b]Correlation of freshman year GPA with Sem 8 GPA; separate Sem 1, Sem 2 GPA not available.

[c]Test negatively weighted in this and subsequent composites.

[d]Sem 1 GPA, SAT-M negatively weighted.

[e]Sem 1 negatively weighted.

[f]Sem 3 GPA negatively weighted.

[g]Both freshman-year GPA and SAT-M negatively weighted.

Table 10. Results of Special Reanalyses Indicating the Slight Contribution of Admissions Measures to Predictions of GPA in Successive Semesters After Taking the Previous Semester GPA Record Into Account

			Correlation of Predictors with GPA by Semester								
Setting	*N*	*Predictors*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*Study*[a]
University A (Illinois)	1,600	All prior GPA	–	556	537	610	592	621	625	612	Humphreys (1968)
		Add Admissions	457	576	542	613	594	623	625	613	
		Difference	–	(020)	(015)	(003)	(002)	(002)	(000)	(001)	
Anonymous university B	530	All prior GPA	–	582	650	693	695	704	675	645	Wilson (1978, 1980)
		Add Admissions	538	624	688	710	715	706	682	648	
		Difference	–	(042)	(038)	(017)	(020)	(002)	(007)	(003)	
Selective college	950	All prior GPA	–	N.A.	692	712	719	719	716	660	Wilson (1978, 1981)
		Add Admissions	564	N.A.	705	713	724	720	720	667	
		Difference	–	N.A.	(013)	(001)	(005)	(001)	(004)	(007)	

Note: Samples involved include only individuals who persisted through the senior year. However some GPAs were missing. In computing the tabled coefficients (decimal omitted) from intercorrelation matrices generated by the referenced studies, the matrices were treated as though the (representative) Ns reported herein were constant.

For Sem 1 GPA the coefficient tabled is for the admissions battery, optimally weighted; for Sem 2 GPA the initial coefficient is the correlation with Sem 1 GPA and the second is the multiple correlation with Sem 1 and the admissions battery as predictors; for Sem 3 GPA the multiple correlations with Sem 1 and 2 GPA, and with the admissions battery added, are shown; etc.

[a]The coefficients in the table were calculated from correlation matrices generated by the referenced studies. The admissions variables varied across studies: ACT Composite and HS Rank for University A; SAT-V, SAT-M, and Rank for University B; SAT-V, SAT-M, Rank, and CB Achievement Test Average for the Selective College. For the latter sample, only Year 1 GPA was available, not separate GPA for Semesters 1 and 2.

- It is relevant to note that the freshman-level test scores did not improve prediction of graduate GPA when added to the cumulative undergraduate GPA. One can only speculate that a similar outcome would have been found for a graduate-level test. The most noteworthy aspect of their findings was the unusually high relationship between the cumulative undergraduate GPA and the graduate GPA ($r = .61$ in the total sample). Coefficients in the neighborhood of .3 are typical for the undergraduate-GPA/graduate-GPA relationship in graduate departmental samples that are *heterogeneous* with respect to undergraduate origin (Wilson, 1982).

Too much should not be made of these dated and isolated findings based on a single sample in which the entire undergraduate and graduate GPA record was earned at the same institution. However, in conjunction with other evidence that has been reviewed, these findings suggest the intriguing hypothesis that in circumstances involving both curricular and institutional continuity, GPA in the first graduate year might be predicted as well from the previous undergraduate record alone as by a best-weighted combination of undergraduate grades and senior-level test scores.

Supplementary Analyses

The findings of the studies reviewed above, concerned directly with predicting post-freshman GPA from freshman (or other early) GPA and admissions measures, are extended by the results of special regression analyses based on the Illinois, anonymous university, and selective college matrices. These analyses were designed to assess the contribution of scores on admissions measures to prediction of a fixed, remote GPA criterion (Semester 8 GPA) and to prediction of GPA in each of eight successive semesters after taking into account earlier semester GPA variables.

Two series of step-wise regression analyses were conducted. In the first, Semester 8 GPA was used as a fixed criterion. The initial set of predictors included Semester 1 GPA, followed by scores on the admissions battery; the second set included GPA in Semesters 1 and 2, followed by admissions scores, etc.; through the last set, which included GPA in Semesters 1 through 7, followed by admissions scores. In the second series, the GPA for each of the eight semesters was treated as a separate criterion variable. GPA and admissions scores were introduced as predictors in the same way.

- Table 9 shows results of the analysis in which Semester 8 GPA was the dependent variable. It is evident (a) that prediction of Semester 8 GPA from the previous GPA record improved as grades were accumulated from the first through the seventh semester and (b) that the admissions battery had very limited incremental validity, especially after the freshman year.
- Table 10 summarizes findings of the analysis in which the criteria were GPA in successive semesters. Using the same pattern of predictors, it is evident that regardless of the semester GPA criterion used, the results were similar: the admissions battery added little to prediction after taking the previous grade record into account.

It is of interest to note that there were some changes in the relative contribution of the respective admissions measures to prediction as GPA in successive semesters was introduced as a criterion. In the Illinois analysis, school rank was more heavily weighted relative to ACT scores in the senior year than in the freshman year. In the anonymous university analysis, school rank was *less* heavily weighted relative to SAT scores in the last two than in the first two semesters, and this pattern tended to hold for the selective college data as well. These results clearly do not provide a basis for generalizations, and alternative interpretive rationales are not immediately evident.[16]

Discussion. From the foregoing, it is reasonable to hypothesize that admissions measures (standard tests and the high school record, separately and in best-weighted combination) are likely to have little incremental validity when combined with early college grades to predict later grades (or other general long-term performance criteria, such as those examined by Willingham, 1962). This is particularly noteworthy when contrasted with consistent evidence that admissions tests do have useful incremental validity when combined with the high school GPA for the purpose of predicting freshman GPA (or later GPA) in college.

Why should tests supplement the high school GPA in predicting college GPA while neither tests nor the high school record provide useful additional information when combined with early college GPA for predicting later GPA? The following line of reasoning is relevant:

- Admissions tests such as the SAT or the ACTP battery have been characterized as representing standard work samples of basic cognitive skills

[16] In a study involving replicated analyses at each of 23 law schools, Powers (1982) examined the relationship of the undergraduate GPA and scores on the Law School Admission Test (LSAT) to independently computed GPA for Year 1, Year 2, and Year 3. Generally speaking, the predictive power of both undergraduate GPA and LSAT decreased from Year 1 through Year 3. In the majority of schools, however, the undergraduate GPA made its greatest relative contribution to prediction of later GPA, while LSAT scores made their greatest relative contribution to prediction of Year 1 GPA. Powers concluded that the results "... were consistent with the test sponsor's interpretation of the LSAT as a measure of the ability to confront new situations and new problems such as those encountered in the first year of law school [as well as] in interpretation of the university GPA as a measure not only of academic competence but also of other qualities (such as persistence and motivation) that are needed over the long haul for the study of law."

called for in the performance of academic tasks.[17] Thus, the correlation between standardized tests and GPA (a much more complex measure) may be explained in terms of a common-elements theory.

- It is reasonable to infer (a) that scores on these standardized work samples help to adjust the high school record, at least in part, for differences among secondary schools in grading standards, and (b) that they supplement the high school GPA for predicting college GPA *primarily because they help to correct for differences in school standards, not because they add unique GPA-criterion-related variance in a composite predictor of college GPA.*
- All the variance in the adjusted high school record (predicted college GPA based on a composite of standard test scores and the high school record) that is useful for predicting post-freshman-year GPA (later GPA) is reflected completely in the observed freshman GPA (earlier GPA). Thus, neither tests nor preadmission composites appear to have significant incremental validity when combined with earlier GPA to predict later GPA within the same institutional context.

A general corollary hypothesis is that the prediction of later grades (in successive marking periods or cumulatively) from all earlier grades is likely to remain relatively stable over time and will not be improved by using standard cognitive test scores under conditions of curricular and programmatic (curricular) continuity. However, tests *are* likely to be useful supplements to earlier GPA in circumstances characterized by institutional and/or curricular discontinuity—as in the typical selective admissions context.

The findings reviewed in this section suggest a substantial level of generality for Willingham's (1962, 1964) conclusion that preadmissions measures are useful because they contribute directly to the primary screening process (i.e., to the initial selection of students). However, in the continuous postadmission secondary secreening process involving at least implicit predictions about the future performance of students (as in setting minimum performance standards for dismissal, for example), preadmissions measures do not make a contribution to the prediction of (or inferences regarding) later performance after taking into account a student's early performance record—that is, soon after admission, college grades may have tapped most of the variance in preadmissions measures that is useful for the purpose of predicting post-freshman-year performance.

[17]Cronbach (1977, p. 2) observes: "It can be seen that the SAT is very nearly a work sample of the interpretation and problem solving required in college, leaving out knowledge of particular subject matter. A person who does poorly on these tests is likely to have similar difficulty with college assignments." ACTP (1973, p. 19) asserts: "At the high-school-to-college transition, work samples simulating activities required of the college students are used in the ACT assessment."

RECAPITULATION

The research findings that have been reviewed are believed to provide substantial support for the several generalizations or working propositions outlined at the outset. These generalizations and key themes in the supporting research evidence are summarized briefly below:

1. Admissions measures (commonly used standard ability tests and the high school record) are essentially as valid for predicting the overall, long-term cumulative grade record as for predicting the short-term, freshman GPA criterion traditionally used in basic predictive validity studies.

This proposition is supported by (a) the generally consistent findings of a limited number of studies employing both cumulative GPA and freshman GPA, as well as by (b) the fact that freshman GPA variance is a continuing element in any longer-term cumulative GPA. Freshman GPA/cumulative GPA correlations higher than .8 have been reported in a variety of different contexts and this level of correlation may be typical. Given such a high and statistically assured relationship between the cumulative GPA and the freshman GPA, it is reasonable to expect that admissions measures that predict the freshman GPA will predict the cumulative GPA at a generally comparable level.

2. Admissions measures (both tests and the high school record) tend to be less valid for predicting independently computed post-freshman-year GPA than for predicting either the cumulative or freshman GPA; generally speaking, there is a gradual decline in their predictive validity as GPAs in successive marking periods (years or terms) are introduced as criteria from the freshman through the senior year.

Because there have been so few four-year longitudinal studies using both independent GPA in successive marking periods and cumulative GPA, the amount of direct empirical support for this proposition is limited. However, the available empirical evidence is reinforced strongly by a conceptual framework provided by Humphreys (1960) that, in effect, anticipates a declining validity gradient for admissions measures from the freshman through the senior year.

- Humphreys (1960) noted that intercorrelation matrices for undergraduate GPA in eight semesters, like other matrices for repeated measures in time series under conditions of growth, learning, or maturation (e.g., repeated measures of mental age, performance on successive shorthand tests during learning, etc.) conform generally to the simplex pattern. At all points in such series, the correlation between a given measurement and the later measurements tends to decrease regularly as the time interval between measurements increases. Given the simplex pattern of correlations for successive GPA, over

time it is reasonable to expect that the correlation of preadmissions measures with these criteria will also follow a declining gradient.

- In a series of empirical studies between 1968 and 1977, involving the prediction of undergraduate GPA by semester, Humphreys and his associates provided evidence that the declining validity gradient was not due to statistical artifacts—that is, such gradients were observed in samples of graduating seniors (controlled experimentally for restriction in range due to attrition among low-scoring students on the admissions variables), in which the standard deviations of GPA variables and GPA reliabilities were similar across semesters. Moreover, college-senior-level scores on the Graduate Record Examinations were found to be more highly correlated with college freshman than with college-senior GPA.
- These findings, along with evidence of four-year stability coefficients exceeding .8 for standard admissions tests, and a simplex pattern of correlation among GPA variables, tend to rule out changes in the rank ordering of students in the general abilities measured by admissions tests as the sole explanation of the declining validity gradient for test scores, and lend support to the hypotheses of gradual change in the nature of the GPA criterion as a partially explanatory factor.

3. The freshman GPA tends to be more valid than admissions measures for predicting post-freshman-year GPA, but shows a similar declining correlational gradient with successive post-freshman GPA. However, the level of correlation between the general cumulative GPA at the end of any given term and (a) GPA in the immediately following term, as well as (b) cumulative GPA for all subsequent terms tends to remain relatively stable from the freshman through senior years.

Empirical evidence for this proposition is quite limited. Very few investigators have been interested in using the freshman GPA to predict later undergraduate GPA or, more generally, in using earlier GPA to predict later GPA within the same setting.

- Studies dating from 1925, including those in the simplex series by Humphreys and associates, have provided evidence that the freshman GPA is more valid than admissions measures for predicting later GPA but is characterized by a similar declining validity gradient.
- However, Juola's (1966) data indicated that the level of correlation between the cumulative GPA at the end of successive quarters and GPA in the immediately following quarter remained relatively stable over the first seven quarters, and Willingham (1962) found a generally flat validity gradient when using the cumulative GPA at the end of successive quarters to predict the cumulative GPA over all subsequent quarters.
- The line of inquiry followed by Willingham and Juola was extended by a special reanalysis of three-semester-GPA correlation matrices, generated in studies not directly concerned with using prior cumulative GPA for predicting later GPA. From these analyses, it was determined that GPA in successive semesters was predicted best by the prior cumulative GPA, that GPA in the immediately previous semester afforded the second-best prediction, and that prediction of a later independent GPA from the earlier GPA remained relatively stable from the freshman through the senior years. The freshman GPA was a better predictor of successive subsequent semester GPA than was Semester 1 GPA, and both in turn tended to be better predictors than scores on the admissions batteries involved; these three predictors had generally declining validity gradients.

4. The limited information that is available suggests, consistently, that commonly used admissions measures have little or no incremental validity when combined with freshman GPA (earlier GPA) for predicting post-freshman GPA (later GPA) within the same institutional setting. The best prediction of GPA in any given marking period is provided by the previous cumulative GPA alone.

Empirical evidence for this proposition is provided by the quite consistent findings of a small number of studies in which investigators were directly concerned with the regression of one or more post-freshman-year GPA criteria on a combination of predictors that included both scores on various admissions measures and the freshman (or other early) GPA. These findings are reinforced by the results of special reanalyses of three correlation matrices that were designed to extend this line of inquiry. Generally speaking, all available evidence indicates little or no improvement in multiple correlation when scores on admissions measures are added to the freshman or other earlier GPA for predicting later GPA within a given setting. The consistency of such findings reinforces Willingham's conclusion (1962, 1964) that soon after admission, college grades may have tapped most of the variance in admissions measures that is useful for the purpose of predicting post-freshman-year performance.

IMPLICATIONS FOR VALIDATION RESEARCH

What are the implications of the findings reviewed here for the traditional practice of using only the freshman GPA in admissions-related predictive validity studies and for validation research generally?

Implications for the Use of the Freshman GPA in Validity Studies

The available evidence lends credibility to the traditional practice of relying upon the first-year or freshman GPA as the criterion in basic admissions-related predictive validity studies.

- It is reasonable to argue on both pragmatic and conceptual grounds (a) that the first-year GPA is a satisfactory surrogate for the overall cumulative GPA , the most comprehensive *GPA* criterion available for use in admissions-related predictive validity studies, and that the two may be thought of as statistically interchangeable. Accordingly, answers to questions regarding the level of validity of admissions measures, separately and collectively, and their relative contribution to prediction, that have been based on numerous first-year studies are likely to be applicable with regard to the prediction of the cumulative average.
- The freshman GPA (or the variable-base overall cumulative GPA) is the only GPA criterion both applicable to and available for essentially all members of each entering cohort, and no post-freshman-year GPA criterion would seem logically to have superior merit as an admissions-related criterion, given the logical dependence of later performance on early performance during the undergraduate years. Post-freshman-year GPA is applicable only to individuals who survive the freshman year and these, almost inevitably, will be a biased sample of the entering cohort, having been screened on all variables associated with persistence, particularly prior grades.
- First-year performance is critical, since students must meet minimum GPA requirements during the first year (and usually for each quarter or semester during that year) *and* cumulatively in order to avoid academic dismissal or probation and to advance in good standing toward program completion. The best predictor of GPA in any marking period is the cumulative GPA prior to that period, and freshman GPA/cumulative GPA correlations above .8 probably are typical.

Thus, from the point of view of selective admission, or primary screening, given the continuous postadmission (secondary) screening to which students are subjected, based on grades during each marking period, the freshman GPA would appear to be the logical target for admissions-related predictions *and* predictive validity studies. Commonly used admissions measures, both the high school record and standard test, should be judged primarily in terms of their validity for predicting that target (or the cumulative GPA) and not in terms of their lower validity for predicting independently computed GPA after the freshman year.

On balance, it appears (a) that the declining validity of admissions measures with respect to GPA in successive undergraduate terms or years reflects a corresponding decline in the relationship between freshman and later grades, and (b) that the major contribution of admissions measures (both the high school record and scores on standardized admissions tests) is to facilitate the primary screening of prospective students—their contribution to the prediction of post-freshman-year performance appears to be indirect, through the post-freshman-year-related GPA variance that they have in common with the freshman-year GPA.

The foregoing should not be taken to mean that it is not important to conduct studies designed to assess the comparative validity of admissions measures for predicting freshman-year and overall cumulative GPA criteria—questions regarding the validity of particular admissions measures for predicting particular criteria in particular admissions contexts require empirical answers. The proposition of "statistical interchangeability" of freshman and cumulative GPA criteria should be viewed as a working hypothesis calling for confirmatory studies.

As suggested elsewhere (Wilson, 1980), in most undergraduate settings the most appropriate cumulative GPA criterion for such comparative studies would be the final (variable-base) cumulative, based on all work completed by students in a given entering cohort, regardless of amount, during their undergraduate careers; in settings characterized by very low attrition rates, the fixed-base four-year cumulative may be appropriate.

General Implications for Validation Research

The declining predictive validity gradient for admissions measures with respect to GPA in successive marking periods from the freshman through the senior year appears to reflect a corresponding decline in the relationship between freshman and later undergraduate grades. More generally, such a declining validity gradient is consistent with expectation, given the tendency noted by Humphreys for matrices of GPA intercorrelations during the undergraduate years to conform generally to the simplex pattern.

This phenomenon has been observed for repeated measures in time series under conditions of growth, learning, or maturation. More empirical evidence is needed to establish the degree of generality of the declining validity and simplex phenomena—for example, based on multiinstitutional data, including correlations among GPA in successive terms and correlations of admissions measures with these GPA (and the cumulative GPA), from a representative array of undergraduate settings. In the analysis of such data, it is important to examine validity gradients (and GPA intercorrelations) in samples of college seniors or graduates classified by field of study as well as in general samples of successful students. A well-designed multiinstitutional study would permit, among other things, (a) the use of large within-major correlation matrices, based on data pooled across in-

stitutions, in assessing, comparatively, college freshman-senior-level validity gradients for the high school record and scores on standardized admission tests, (b) analysis of trends across institutions in patterns of GPA intercorrelations between the freshman and the senior year, and (c) assessment of the contribution of admissions measures to predicting post-freshman-year GPA after taking the freshman GPA into account.

Based on the available evidence, there are significant changes in rank ordering among successful students in terms of GPA between the freshman and the senior year. Evidence stemming from the work of Humphreys and his associates, within the framework of the simplex model, indicates that it is important to consider both the possibility of "changes in people" and "changes in the factorial composition of the GPA criterion" in any general explanatory rationale for such rank order changes and the associated declining validity phenomenon. Research is needed to explore these hypotheses and to extend and elaborate lines of inquiry initiated by Humphreys et al. For example, more definitive evidence is needed regarding the stability of scores on standard admissions tests between the high-school-senior and college-senior level and similarity in patterns of correlation for high-school-senior and college-senior level scores on the same or similar tests with respect to undergraduate GPA in successive marking periods.

Research is needed that is concerned with the identification of college-senior-level measures of cognitive or noncognitive characteristics that might correlate more highly with senior GPA than with freshman GPA in samples controlled for field of study, as is similar research concerned with the identification of preadmissions or freshman-level measures that might have incremental validity for predicting senior GPA after taking freshman or other earlier GPA into account, again in samples of college seniors classified by major field.

Available evidence suggests the hypothesis that scores on standard tests of developed verbal and quantitative reasoning abilities, such as those used in college admission, are likely to have incremental value when combined with earlier GPA for predicting later GPA only in circumstances (a) in which the relationship between the earlier GPA and the later GPA variables is attenuated by heterogeneity of institutional origin with respect to the earlier GPA and attendant differences in grading standards or (b) in which earlier and later grades relect performance in substantially different learning tasks (curricular areas).

This hypothesis derives from evidence (a) that tests supplement the high school GPA for predicting college GPA but do not supplement the freshman, or other early GPA in the prediction of later GPA within the same setting; (b) that within a given undergraduate setting, the best prediction of GPA in a given marking period is provided by the cumulative GPA prior to that period; and (c) that the correlation of cumulative GPA at the end of a given term with GPA in the next term, or the cumulative GPA over all subsequent terms, remains relatively stable from the freshman through the senior year.

- Research is needed to determine whether this general pattern of relationships will obtain for GPA data extending from, say, the undergraduate freshman year into graduate study, in graduate-level samples controlled for both undergraduate origin and field of study, as suggested by the work of Weber, Brink, and Gilliland (1942), who employed both undergraduate cumulative GPA and test data for predicting graduate GPA in a sample composed of students whose undergraduate and graduate study had been completed at the same university and in the same field of study.

It is important to recognize that research such as that suggested above is more concerned with illuminating possible changes in what is being measured by grades and tests between the freshman and the senior year in college, and with hypothesis testing, than with accuracy of prediction, per se. And, unlike traditional short-term first-year studies, the research is not likely to contribute immediately or directly to the implementation of a recurring, pragmatic function such as selective admission. Moreover, studies concerned with the analysis of successive GPA criteria spanning several years inevitably pose much more complex problems of design, data collection, analysis, and interpretation than traditional studies concerned primarily with the prediction of a single, fixed criterion. Given these considerations, it will be tempting for those concerned with selective admission over the next 50 years to remain as disinterested in post-freshman-year research as their counterparts over the past 50 years were, judging from the limited amount of research unearthed in the present review.

However, testing-program-sponsored validity study services have been instrumental in encouraging the widespread and recurring participation of test users in traditional validity studies, and testing agencies can play a similar role in connection with post-freshman-year research.

- In order to promote the participation of undergraduate institutions in post-freshman-year studies, testing agencies could not only (a) extend their standard data collection, analysis, and reporting routines to encompass the treatment of independently computed post-freshman-year GPA, especially the senior-year GPA, as criteria, but also (b) actively encourage participation in their validity study services to provide post-freshman-year criterion data periodically, albeit perhaps less frequently than they provide the traditional first-year GPA criterion.
- In addition to the foregoing, which would result in more attention to post-freshman-year validity research, a more immediate contribution of undergraduate testing agencies would be to sponsor ad

hoc, multiinstitutional studies along lines suggested earlier.[18]

Such steps by testing agencies could have the important incidental effect of encouraging institutions of higher education to develop computerized records systems that are oriented to the collation of admissions data (on the entry characteristics of students) with data on student progress (such as GPA computed term by term and cumulatively, highest level of progress attained, and choice of major field) to form cohort files. As noted elsewhere (Wilson, 1980), admissions records that contain precollege information and records on student progress (registration and transcript records and related systems) frequently are maintained separately without routines for merging into individual record files for each cohort. This reflects the traditional lack of institutional orientation to routine, systematic analysis of the experience of subgroups of students in successive cohorts—subgroups defined in terms of level of ability, sex, race, and/or other important personal and background characteristics.

Better systems for student flow analysis would enable colleges and universities not only to generate the more complex patterns of data required for the limited purpose of assessing trends in GPA intercorrelations and validity gradients from the freshman through the senior year, but also, more generally, to obtain regularly empirical evidence regarding the relationship of student characteristics to important educational outcomes—evidence needed for informed assessment, evaluation, and policy review on matters affecting the admission, retention, and educational development of students.

REFERENCES

Alderman, D. L. 1981. *Measurement Error and SAT Score Change* (College Board Report No. 81-9). New York: College Entrance Examination Board.

American College Testing Program. 1973. *Assessing Students on the Way to College* (Technical Report for the ACT Assessment Program Vol. 1). Iowa City, IA: ACT.

Anderson, J. E. and Spencer, L. T. 1926. The Predictive Value of the Yale Classification Tests. *School and Society* 24, 305-312.

Breland, H. M. 1978. *Population Validity and College Entrance Measures* (RB-78-19). Princeton, NJ: Educational Testing Service.

Brush, E. N. 1941. Mechanical Ability as a Factor in Engineering Aptitude. *Journal of Applied Psychology* 25, 300-312 (cited in French, 1957).

Carlson, A. B., and Werts, C. E. 1976. *Relationships among Law School Predictors, Law School Performance, and Bar Examination Results* (PR-76-26). Princeton, NJ: Educational Testing Service.

Chase, E. B. 1960. A Study of Undergraduate Records of Graduates from Hunter College Who Later Earned Doctorates. *Journal of Experimental Education* 29(1), 49-60.

Creech, F. R., Freeberg, N. E., Rock, D. A., Wilson, K. M., and Young, K. 1977. *Comparative Analysis of Postsecondary Occupational and Educational Outcomes for the High School Class of 1972* (Project Report 77-7). Princeton, NJ: Educational Testing Service.

Cronbach, L. J. 1977. *Educational Psychology*. New York: Harcourt Brace Jovanovich.

Davis, J. A., and Frederiksen, N. 1954. *Academic Performance of Public and Private School Graduates at Princeton* (RB-54-21). Princeton, NJ: Educational Testing Service.

Dumont, R. G., and Troelstrup, R. L. 1981. Measures and Predictors of Educational Growth with Four Years of College. *Research in Higher Education* 14, 31-47.

Dvorak, A., and Langen, T. 1962. The Washington Program for Differential Guidance. *Journal of College Student Personnel* 4, 61–78.

Dwyer, P. S., et al. 1940. *A Statistical Summary of the Records of Students Entering the University of Michigan as Freshmen in the Decade 1927-1936* (Administrative Studies, I(4). Ann Arbor: University of Michigan (cited in French, 1957).

Etaugh, A. F., Etaugh, C. F., and Hurd, D. E. 1972. Reliability of College Grades and Grade Point Averages: Some Implications for Prediction of Academic Performance. *Educational and Psychological Measurement* 32, 1045-1050.

Farver, A. S., Sedlacek, W. E., and Brooks, G. C. 1975. Longitudinal Predictions of University Grades for Blacks and Whites. *Measurement and Evaluation in Guidance* 7, 243-250.

Fincher, C. 1974. Is the SAT Worth Its Salt? An Evaluation of the Scholastic Aptitude Test in the University System of Georgia Over a Thirteen-Year Period. *Review of Educational Research* 44(3), 293-305.

Fishman, J. A. 1963. Some Social-Psychological Theory for Selecting and Guiding College Students. In N. Sanford, ed., *The American College*. New York: Wiley.

Fishman, J. A., and Pasanella, A. 1960. College Admissions Selection Studies. *Review of Educational Research* 30, 298-310.

Frederiksen, N. O. Unpublished study (cited in French, 1957).

Freeman, F. S. 1931. Predicting Academic Survival. *Journal of Educational Research* 23, 113-123.

French, J. W. 1957. *Validation of the SAT and New Item Types Against Four-Year Academic Criteria* (RB-57-4). Princeton, NJ: Educational Testing Service.

French, J. W. 1958. Validation of New Item Types Against Four-Year Academic Criteria. *Journal of Educational Psychology* 49, 67-76.

Gough, G. H., Hall, W. B., and Harris, R. E. 1964. Evaluation of Performance in Medical Training. *Journal of Medical Education* 39, 679-692.

[18] See Powers (1982), for an example of such a study involving replicated analyses of relationships of undergraduate GPA and the Law School Admission Test of law school GPA in Years 1, 2, and 3.

Hardesty, L. 1980. *Use of Multiple Regression to Predict Academic Acheivement at a Small Liberal Arts College* (ERIC Document Reproduction No. ED 185 960). Greencastle, IN: DePauw University.

Hills, J. R., Bush, M. L., and Klock, J. A. 1964. Predicting Grades Beyond the Freshman Year. *College Board Review* 54, 23-25.

Horst, P. 1959. Differential Prediction of Academic Success. In T. R. McConnell, ed., *Selection and Educational Differentiation*. Berkeley: University of California, Center for the Study of Higher Education.

Humphreys, L. G. 1960. Investigations of the Simplex. *Psychometrika* 25(4), 313-323.

Humphreys, L. G. 1967. Problems in Personnel Research. In A. L. Fortuna, ed., *Personnel Research and Systems Advancement*. Springfield, VA: Clearinghouse for Federal Scientific and Technical Information, Sills Bldg., 5285 Port Royal Rd.

Humphreys, L. G. 1968. The Fleeting Nature of the Prediction of College Academic Success. *Journal of Educational Psychology* 59(5), 375-380.

Humphreys, L. G. 1976a. Prediction of Grades Is Fleeting: A Comment. *Journal of Educational Psychology* 68, 518-519.

Humphreys, L. G. 1976b. The Phenomena Are Ubiquitous–but the Investigator Must Look. *Journal of Educational Psychology* 68, 521.

Humphreys, L. G., Levy, J., and Taber, T. 1973. Predictability of Academic Grades for Students of High and Low Academic Promise. *Educational and Psychological Measurement* 33, 385-392.

Humphreys, L. G., and Taber, T. 1973. Postdiction Study of the Graduate Record Examination and Eight Semesters of College Grades. *Journal of Educational Measurement* 10(3), 179-184.

Jones, M. H., and Case, H. W. 1955. The Validation of a New Aptitude Examination for Engineering Students. *Educational and Psychological Measurement* 15(4), 502-508.

Juola, A. E. 1966. Prediction of Successive Terms Performance in College from Tests and Grades. *American Educational Research Journal* 3, 191-197.

Kallingal, A. 1971. The Prediction of Grades for Black and White Students at Michigan State University. *Journal of Educational Measurement* 8(4), 263-265.

Langen, T. D. F. 1965. *An Investigation of Additional Predictor and Criterion Variables for the Washington Precollege Testing Program with Subdivision by Sex and Extent of Achievement* (Doctoral dissertation, University of Washington). (University Microfilms No. 65-11, 471).

Lewis, J. W. 1964. Precollege Variables as Predictors of Freshman, Sophomore, and Junior Achievement. *Educational and Psychological Measurement* 24(2), 353-356.

Lewis, W. A. 1966. Early Prediction of College GPA Using Precollege School Grades. *Journal of Educational Measurement* 3(1), 35-36.

Leyman, L. 1967. Prediction of Freshman and Sophomore Grade Point Averages of Women Physical Education Major Students. *Educational and Psychological Measurement* 27, 1139-1141.

Lin, P.-C., and Humphreys, L. G. 1977. Predictions of Academic Performance in Graduate and Professional School. *Applied Psychological Measurement* 1(2), 249-257.

Lins, L. J., and Abell, A. P. 1965. *Attendance Patterns of Fall 1958 New Freshman for Twelve Semesters After Entrance*. Madison: University of Wisconsin, Office of Institutional Studies.

Loyd, B. H., Forsyth, R., and Hoover, H. D. 1981. Relationship of Elementary and Secondary School Achievement Test Scores to Later Academic Success. *Educational and Psychological Measurement* 41, 1117-1124.

Lunneborg, C. E. 1966. A Research Review of the Washington Precollege Testing Program. *Journal of Educational Measurement* 3, 157-166.

Lunneborg, C. E., and Lunneborg, P. W. 1970. Relations between Aptitude Changes and Academic Success During College. *Journal of Educational Psychology* 61, 169-173.

Lunneborg, P. W. 1977. Longitudinal Criteria of Predicted College Achievement. *Measurement and Evaluation in Guidance* 9(4), 212-213.

Mauger, P. A. and Kolmodin, C. A. 1975. Long-term Predictive Validity of the Scholastic Aptitude Test. *Journal of Educational Psychology* 67(6), 847-851.

McConnell, T. R. 1934. Change in Scores on the Psychological Examination of the American Council on Education from Freshman to Senior Year. *Journal of Educational Psychology* 25, 66-69.

Miner, J. G. 1927. Scholarship and Intelligence: Relationships for the Same Groups throughout the College Course. *Personnel Journal* 6, 113-118.

Nichols, R. C. 1964. Effects of Various College Characteristics on Student Aptitude Test Scores. *Journal of Educational Psychology* 55(1), 45-54.

Olsen, M., and Schrader, W. B. 1959. *The Use of Preliminary and Final Scholastic Aptitude Test Scores in Predicting College Grades* (SR-59-19). Princeton, NJ: Educational Testing Service.

Powers, D. E. 1982. Long-term Predictive and Construct Validity of Two Traditional Predictors of Law School Performance. *Journal of Educational Psychology* 74, 568-576.

Price, F. W. and Kim, S. 1976. The Association of College Performance with High School Grades and College Entrance Test Score. *Educational and Psychological Measurement* 36, 965-970.

Prien, E. P., and Botwin, D. E. 1966. The Reliability and Correlates of an Achievement Index. *Educational and Psychological Measurement* 26, 1047–1052.

Ramist, L., Criterion-Related Validity of Tests Used in the College Board's Admissions Testing Program. Unpublished manuscript, Educational Testing Service.

Rock, D., and Werts, C. 1979. *An Analysis of Time-Related Increments and/or Decrements for GRE Repeaters Across Ability and Sex Groups*. GREB No. 77-9R. Princeton, NJ: Educational Testing Service.

Rogers, A. L. 1925. Mental Tests for the Selection of University Students. *British Journal of Psychology* 15, 405-415.

Schell, R. E. 1975. *Semester Indices of Successful Students*

as a Function of the Semester in Attendance (ERIC Document Reproduction No. ED 114 030). Oswego, NY: State University College.

Schoenfeldt, L. F., and Brush, D. H. 1975. Patterns of College Grades Across Curricular Areas: Some Implications for GPA as a Criterion. *American Educational Research Journal* 12(3), 313-321.

Schrader, W. B. 1967. *The Relation of College Board SAT Scores to Graduate Record Examination Scores*. Unpublished manuscript, Educational Testing Service.

Schrader, W. B. 1971. The Predictive Validity of College Board Admissions Tests. In W. H. Angoff, ed., *The College Board Admissions Testing Program: A Technical Report on Research and Development Activities Relating to the Scholastic Aptitude Test and Achievement Tests*. New York: College Entrance Examination Board.

Segel, D. 1934. *Prediction of Success in College: A Handbook for Administrations and Investigators Concerned with the Problems of College Admittance or Guidance of College Students* (Bulletin, No. 15). Washington, DC: U.S. Government Printing Office.

Sgan, M. R. 1964. An Alternative Approach to Scholastic Aptitude Tests as Predictors of Graduation Rank at Selective Colleges. *Educational and Psychological Measurement* 24(2), 347-352.

Siegelman, M. 1971. SAT and High School Average Predictions of Four Year College Achievement. *Educational and Psychological Measurement* 31, 947-950.

Snedecor, G. W. 1965. *Statistical Methods*. Ames: Iowa State College Press.

Spindt, H. A. 1959. Improving the Prediction of Academic Achievement. In T. R. McConnell, ed., *Selection and Educational Differentiation*. Berkeley: University of California, Center for the Study of Higher Education.

University. *Study of Undergraduate Education–A Study of the Variables Used to Determine Freshman Admission.* (Anonymous university, unpublished report April 10, 1956. Cited in French, 1957.

Weber, J., Brink, W. G., and Gilliland, A. R. 1942. Success in the Graduate School Determined by Using Statistical Measures of Prediction. *Journal of Higher Education* 13, 19-24.

Werts, C. E., Linn, R. L., and Jöreskog, K. G. 1978. Reliability of College Grades from Longitudinal Data. *Educational and Psychological Measurement* 38, 89-95.

Willingham, W. W. 1962. *Longitudinal Analysis of Academic Performance* (Research Memorandum 62-5). Atlanta: Georgia Institute of Technology.

Willingham, W. W. 1964. The Interpretation of Relative Achievement. *American Educational Research Journal* 1(2), 101-112.

Wilson, K. M. 1967. *Type of Analyses Underway or Planned: A Selective Review* (Memorandum). Poughkeepsie, NY: Vassar College, College Research Center.

Wilson, K. M. 1970. *Predicting Cognitive Performance in Different Colleges: Toward the Use of an Expected Output Dimension in Educational Evaluation.* Paper presented at the 137th meeting of the American Association for the Advancement of Science, Chicago.

Wilson, K. M. 1976. The Utility of a Standard Composite for Forecasting Academic Performance in Several Liberal Arts Colleges. *Research in Higher Education* 5, 192-213.

Wilson , K. M. 1978. *Predicting the Long-term Performance in College of Minority and Nonminority Students: A Comparative Analysis in Two Collegiate Settings* (RDR 77-78, No. 3 and RB-78-6). Princeton, NJ: Educational Testing Service.

Wilson, K. M. 1980. The Performance of Minority Students Beyond the Freshman Year: Testing a "Late Bloomer" Hypothesis in One State University Setting. *Research in Higher Education* 13, 23-47.

Wilson, K. M. 1981. Analyzing the Long-term Performance of Minority and Nonminority Students: A Tale of Two Studies. *Research in Higher Education* 15, 351-375.

Wilson, K. M. 1982. *A Study of the Validity of the Restructured GRE Aptitude Test for Predicting First-Year Performance in Graduate Study*. GREB No. 78-6R (ETS RR82-34).

Appendix. Intercorrelations of Preadmissions Variables and Undergraduate Semester GPA for Samples of Graduating and Senior-Level Students from the University of Illinois, Anonymous State University, and Selective Liberal Arts College

Original data matrix: University of Illinois

	HS RK	ACT COMP	SEM1 GPA	SEM2 GPA	SEM3 GPA	SEM4 GPA	SEM5 GPA	SEM6 GPA	SEM7 GPA	SEM8 GPA
HS RK	1.0000	0.3930	0.3870	0.3410	0.2780	0.2700	0.2400	0.2560	0.2400	0.2220
ACT COMP	0.3930	1.0000	0.3750	0.2980	0.2370	0.2550	0.2380	0.2520	0.2190	0.1730
SEM1 GPA	0.3870	0.3750	1.0000	0.5560	0.4560	0.4390	0.3990	0.4150	0.3870	0.3420
SEM2 GPA	0.3410	0.2980	0.5560	1.0000	0.4900	0.4450	0.4180	0.3830	0.3640	0.3390
SEM3 GPA	0.2780	0.2370	0.4560	0.4900	1.0000	0.5620	0.4960	0.4560	0.4450	0.3540
SEM4 GPA	0.2700	0.2550	0.4390	0.4450	0.5620	1.0000	0.5120	0.4690	0.4420	0.4160
SEM5 GPA	0.2400	0.2380	0.3990	0.4180	0.4960	0.5120	1.0000	0.5510	0.5000	0.4530
SEM6 GPA	0.2560	0.2520	0.4150	0.3830	0.4560	0.4690	0.5510	1.0000	0.5440	0.4820
SEM7 GPA	0.2400	0.2190	0.3870	0.3640	0.4450	0.4420	0.5000	0.5440	1.0000	0.5410
SEM8 GPA	0.2220	0.1730	0.3420	0.3390	0.3540	0.4160	0.4530	0.4820	0.5410	1.0000

N = 1,600

Source: Humphreys (1968)

Original data matrix: Anonymous State University

	SAT-V	SAT-M	HS RK	S1 GPA	SM2 GPA	SM3 GPA	SM4 GPA	SM5 GPA	SM6 GPA	SM7 GPA	SM8 GPA
SAT-V	1.0000	0.6220	0.5150	0.4030	0.3990	0.4460	0.4640	0.4720	0.4000	0.3530	0.3550
SAT-M	0.6220	1.0000	0.4660	0.4110	0.3900	0.4740	0.4460	0.4870	0.3630	0.2650	0.3170
HS RK	0.5150	0.4660	1.0000	0.4910	0.4450	0.4400	0.4030	0.4380	0.3810	0.3200	0.3340
S1 GPA	0.4030	0.4110	0.4910	1.0000	0.5820	0.5370	0.4950	0.4910	0.4570	0.3700	0.3450
SM2 GPA	0.3990	0.3900	0.4450	0.5820	1.0000	0.6110	0.5320	0.5180	0.4930	0.3900	0.3900
SM3 GPA	0.4460	0.4740	0.4400	0.5370	0.6110	1.0000	0.6660	0.5960	0.5580	0.4730	0.4410
SM4 GPA	0.4640	0.4460	0.4030	0.4950	0.5320	0.6660	1.0000	0.6350	0.5620	0.4790	0.4640
SM5 GPA	0.4720	0.4870	0.4380	0.4910	0.5180	0.5960	0.6350	1.0000	0.6580	0.6010	0.5420
SM6 GPA	0.4000	0.3630	0.3810	0.4570	0.4930	0.5580	0.5620	0.6580	1.0000	0.6200	0.5510
SM7 GPA	0.3530	0.2650	0.3200	0.3700	0.3900	0.4730	0.4790	0.6010	0.6200	1.0000	0.5640
SM8 GPA	0.3550	0.3170	0.3340	0.3450	0.3900	0.4410	0.4640	0.5420	0.5510	0.5640	1.0000

Typical N = 530

Source: Wilson (1978, 1981)

Original data matrix: Anonymous Selective Liberal Arts College

	SAT-V	SAT-M	HS RK	CB ACH	S1, 2 GPA	SM3 GPA	SM4 GPA	SM5 GPA	SM6 GPA	SM7 GPA	SM8 GPA
SAT-V	1.0000	0.5270	0.2310	0.6770	0.4220	0.3970	0.3370	0.3870	0.3530	0.3490	0.3290
SAT-M	0.5270	1.0000	0.2870	0.6920	0.3750	0.3560	0.2890	0.3240	0.2820	0.2760	0.1950
HS RK	0.2310	0.2870	1.0000	0.3040	0.3580	0.2910	0.2750	0.2660	0.2480	0.1930	0.2000
CB ACH	0.6770	0.6920	0.3040	1.0000	0.5140	0.4490	0.4010	0.4210	0.3890	0.3990	0.3360
S1, 2 GPA	0.4220	0.3750	0.3580	0.5140	1.0000	0.6920	0.6420	0.6280	0.5790	0.5670	0.4630
SM3 GPA	0.3970	0.3560	0.2910	0.4490	0.6920	1.0000	0.6660	0.6440	0.5980	0.5100	0.5120
SM4 GPA	0.3370	0.2890	0.2750	0.4010	0.6420	0.6660	1.0000	0.6290	0.5840	0.5410	0.4930
SM5 GPA	0.3870	0.3240	0.2660	0.4210	0.6280	0.6440	0.6290	1.0000	0.6670	0.6210	0.4970
SM6 GPA	0.3530	0.2820	0.2480	0.3890	0.5790	0.5980	0.5840	0.6670	1.0000	0.6260	0.5500
SM7 GPA	0.3490	0.2760	0.1930	0.3990	0.5670	0.5100	0.5410	0.6210	0.6260	1.0000	0.6060
SM8 GPA	0.3290	0.1950	0.2000	0.3360	0.4630	0.5120	0.4980	0.4970	0.5500	0.6060	1.0000

Typical N = 950

Source: Wilson (1978, 1980)

Note: N/coefficient is more nearly constant in the Illinois than in the state university and selective college samples although only graduating or senior-level students are represented. For purposes of the special reanalyses reported herein, the matrices were treated as though N were constant.